

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG
CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA
TTAAACAAGGCTGGCCTTGTACTTGGAATACTGAGTTGTTTAGGACTTTCTATTGTGGCAAACCT
CCAGAAAACAACCCCTTTTGTGTCACATGTAAGTGAGCTGTGCTTACCTTTGGTATGGGCTCAT
TATATATGTTTGTTCAGACCATCCTTTTCCTACCAAATGCAGCCCAAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTAGCATGCTGACTTGCTC
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTT
GGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA
TGGATTAACCCCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA
GAGATATTTTGATGAAAGGATAAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG
TTCACAGAAGTTGCTTATTCCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIIFSITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC
 GTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCTTC
 CTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTTTG
 GATTTGGGTGGCTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTGAGTATGTTGTACAG
 GTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCATTGTTTGAGCTCATCATCTTTGAAATCTTAGG
 AGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGAATTCTGCTGATCCTGG
 TTTTCATGGTGCCTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGA
 CTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCC
 CATTCTCAGCCCAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTGGTGTGATTGGAG
 TGACTCTCATGGCTCTTCTTTCTGGATTGGTGTGTCAACTGCCCATACACTTACATGTCTTACTTC
 CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT
 CATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAAGTGCATAACA
 AACCATCAGGTTTCTGGGAATGATAAAAAGTGTACCACCTTCAGCATCAGGAAGTGAATCTTACT
 CTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTCTGGAACAGCTGATCT
 ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTCAAGGGGAAATATTTAATTTTCTTGGTT
 ACTTTTTCTCTATTTACTGTGTTTGGAAAATTTTCATGGCTACCATCAATATTGTTTTTGATCGAGTT
 GGGAAAACGGATCCTGTCAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGT
 GAAGTTTTGGTCCCAACACATTTCTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGAT
 TGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCCAATGTCATTGTCCTG
 CTATTAGCACAGATAATGGGCATGTACTTTGTCTCTCTGTGCTGCTGATCCGAATGAGTATGCCTTT
 AGAATACCGCACCATAATCACTGAAGTCTTGGAGAAGTGCAGTTCAACTTCTATCACCGTTGGTTTG
 ATGTGATCTTCCTGGTCAGCGCTCTCTTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACCA
 GAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAATTTA
 GATATAAGAGGGGGGAAAATGGAACAGGGCCTGACATTTTATAAACAACAAAATGCTATGGTAGC
 ATTTTTCACCTTCATAGCATACTCCTTCCCCGTGAGGTGATACTATGACCATGAGTAGCATCAGCCAG
 AACATGAGAGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGG
 TGTAGAGGCGGAGAGGAGCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGT
 CTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTAAAGGTTACATGGAAAAGGTTATAGCTTTG
 CCTTGAGATTGACTCATTAAAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCG
 ACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

FIGURE 4

LNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTMYFFWKLGDP
 FFILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERLLQ
 TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDALAEELSRQ
 LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDVPVTRGIEI
 TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLLAQIMGMY
 FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
 425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC
 AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT
 GTTTCCTGGCTCTGAAGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCCATCCAGGT
 GTCATGCAGAAATTATGGGGATCACCTTTGTGAGCAAAAAGGCGAACCAGCAGCTGAATTTACAG
 AAGCTAAGGAGGCCTGTAGGCTGTGGGACTAAGTTTGGCCGGCAAGGACCAAGTTGAAACAGCC
 TTGAAAGCTAGCTTTGAACTTGCAGCTATGGCTGGGTTGGAGATGGATTCTGTGGTCTCTAG
 GATTAGCCCAACCCCAAGTGTGGGAAAAATGGGGTGGGTGCTCTGATTGGAAGGTCCAGTGA
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAAGTCTGTGCATTCCAGAA
 ATTATCACCACCAAGATCCCATATTTCAACACTCAAAGTCAACACAAACAGAAATTTATTGT
 CAGTGACAGTACCTACTCGGTGGCATCCCTTACTCTACAATACCTGCCCTACTACTACTCCTC
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATGATTTGTGTACAGAAAGTTTATG
 GAACTAGCACCATGTCTACAGAACTGAACCATTTGTTGAAAATAAAGCAGCATTCAGAATGA
 AGCTGTGGGTTTGGAGGTGTCCCCACGGCTCTGCTAGTGCTTGCTCTCTCTCTTTGGTGCTG
 CAGCTGGTCTTGGATTTTGTATGTCAAAAGGTATGTGAAGGCCTTCCCTTTTACAACAAGAAT
 CAGCAGAAGGAAATGATCGAAACCAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCTTAA
 TGAGGAATCAAAGAAAAGTATAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACTACCGTGC
 GATGCCTGGAAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTT
 CATGCTCCTTACCCTGCCCGAGCTGGGGAAATCAAAGGGCCAAAGAACCAGAAAGAAAGTCCA
 CCCTTGGTTTCCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAGAGAAT
 GCCCTTCTCCTTATTGTAACCTGTCTGGATCCTATCCTCCTACCTCCAAGCTTCCACGGCCT
 TTCTAGCCTGGCTATGTCTTAATAATATCCCACTGGGAGAAAGGAGTTTGGCAAAGTGAAGGAC
 CTAAAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGGTGGGTTG
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGACCTTTCTTCA
 GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCTTCTGAGCCCGGTAAGAGCAAAAGAAT
 GGCAGAAAAGTTTAGCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAA
 GCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACTGTGAGCAGGGACTGTAAAC
 ACAGACAGGGTCAAAGTGTCTTCTGTAACACATTGAGTTGGAATCACTGTTTAGAACACACACA
 CTTACTTTTTCTGGTCTCTACCACTGCTGATATTTCTCTAGGAAATATACTTTTACAAGTAACA
 AAAATAAAAACTCTTATAAATTTCTATTTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT
 ACTCAGTAATTTGTTTAAAAAGTAATAAAATTCAACAAACATTTGCTGAATAGCTACTATATGTC
 AAGTGTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATGCACATAGTAG
 AACGCTATCTGGGAAGCTATTTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAACTAAT
 TTTTATTTTGTCTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATATAACCTTAATT
 TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAGCACATTTTAAAGTGCC
 ATTAACAAATGTATCACTAGCCCTCCTTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATT
 TGTGACAAAAAATTAAAGCATTTAGAAAACCTT

FIGURE 6

MARCFSLVLLLSIWTRLLVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGCGCCGCCCCACCGCGCGCTCCCGCATCTGCACCCGAGCCCGGC
GGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGCGCCGCGAGCGCAACTCGGTCCAGTCCGGGCGG
CGGCTGCGGGCGCAGAGCGGAGATGCAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGCTGGCGG
CGGCGGTCCCCACGCCCCCGCGCCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCCG
GCTCTCAGCTACCCGCGAGGAGGCGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAAGTATGAT
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAAGTGAACCTGGCAAACCTTACCTCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACAGAC
TGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCCAGCATGTAAGTCCAGTTTGCCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGTGTGGAGACCA
GCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCAGGGGCGAGCAATGGGACCATCTGTGACA
ACCAGAGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCTGCTGTTCCCTGTGTGC
ACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCTCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG
CCGCGCTGCACCTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA
TAGAAATAGCTAATTTATTTCCCGAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCCTA
CATCTTCTTCCAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTAGCT
CCCCCAGGCTGTTCTCCAGGCTTACAGTCTGGTGTGCTGGGAGAGTCAGGCAGGTTAACTGCA
GGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTGCCTCTACCAGTTGGCAGACAGCCG
TTTGTCTACATGGCTTTGATAATTGTTTGGAGGGAGGAGATGGAACAATGTGGAGTCTCCCTC
TGATTGGTTTTGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAACATCAACCTGGCAAAAATG
CAACAAATGAATTTTCCACGAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCTTACAGTGTGTC
AGATGAATGTTCTGTTACCCCTGCATTACATGTGTTTATTCATCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTGTGACTCTAAGCTCAGTGTCTCT
CTCCACTACCCACACAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGATGGGAT
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAATAATTCTCACATCCCTCTAAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGCTCCTCTAATGAAGACAATGAT
ATTGACACTGTCCCTCTTTGGCAGTGCATTAGTAACCTTGAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACCTCTACGCCACTCCACAAATGATG
TTTTCAGGTGTGATGACTGTTGCCACCATGTATTATCCAGAGTTCTTAAAGTTAAAGTTGCA
CATGATTGTATAAGCATGCTTTCTTTGAGTTTAAATTTATGTATAAACATAAGTTGCATTTAGAA
ATCAAGCATAAATCAC'TTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAVPTAPAPAPTATSAPVKPGPALSYQEEATLNEMFREVEELMEDTQHKL
 RSAVEEMEAEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGQMFSE
 TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRLCTRDSECCGDQLCVWGHC
 TKMATRGSNGTICDNQRDCQPLCCAFQRGLLFPVCTPLPVEGELCHDPASRLDLITWELEPDG
 ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
 RSLTEEMALGEFAAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
 327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

[illegible]

APP ID=10063518

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVVLGLFWTL
 NWVLALGQCVLGAFASFYWAFHKPDIPFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
 EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
 IVRVVVLDKVTDLLFFGKLLVVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
 VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCTGCTGCCCCGC
CAGCCGCAACTCCACCGTGAGCCGCCTCATCTTACGTTCTTCTCTTCCCTGGGGGTGCTGGTGTCCA
TCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGTGAGGAGGGGGCC
GGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGGCTCCCTGCTTGGCTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGGCGGCCTTCTTCTTCTTTTACCCTGCTCATGCTCTGCGTGAGCAGCA
GCCGGGACCCCGGGCTGCCATCCAGAATGGGTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTC
ACCGTGGGTGCCTTCTACATCCCTGACGGCTCCTTACCAACATCTGGTTCTACTTCGGCGTCTGGG
CTCCTTCTCTTTCATCTCATCCAGCTGGTGTGCTCATCGACTTTGCGCACTCCTGGAACAGCGGT
GGCTGGGCAAGGCCGAGGAGTGCATTCCCGTGCCTGGTACGCAGGCCTCTTCTTCTTACTCTCTCTC
TTCTACTTGCTGTCGATCGCGGCGTGGCGCTGATGTTTCATGTACTACACTGAGCCAGCGGTGCCA
CGAGGGCAAGTCTTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCATCGCTGCTGTCTCTGC
CCAAGTCCAGGACGCCAGCCAACTCGGGTCTGCTGCAGGCCTCGGTCATCACCTCTACACCATG
TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTTGCCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCGAGCATTG
TGGGCCTCATCATCTTCTCTCTGTCACCTCTTTCATCAGTCTGCGCTCCTCAGACCACCGGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGT
GGCAGCCTGTGAGGGCCGGGCTTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCTTCTTCC
ACTTCTGCCTGGTGTGCGCTCACTGCACGTCATGATGACGCTACCAACTGGTACAAGCCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCGCTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCT
CCTCTACCTGTGGACCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGTTGAGGCAGCCTCA
CAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCCCCCTC
CCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCCCAGCTCCAGGACCTGCCCCTGAGCCGGGC
CTTCTAGTCGTAGTGCCCTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCTCTTCCCTTCCCTCCTCCTGTTGCCATACTCAGCATCTCGGATGAA
AGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAACCTCCACACAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTCTCTGGTACGTCCCCAGGGGACCCTGCCCCCTTCTCTG
GACTTCGTGCCTTACTGAGTCTCTAAGACTTTTCTAATAAACAAGCCAGTGCCTGTAAAAAAA

FIGURE 12

MGACLGACSLSCASCCLGSA PCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEAGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQRWLKAE
ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGVITYSY
SFFHFCLVLASLHVMMTLTNWYKPGETRMISTWTAVVVKICASWAGLLLYLWTLVAPLLLRNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCACCCGTTAAGGTGTCTTCTTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCAGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA
CTCATGGCCAGGATTGAGTCTTATGAAGGAAGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTTGTGTCACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATATTTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGCATATGCTGTGTGCAG
ACTGCGCCATGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG
TGATCCTTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTTCATTCATC
CTTGCCCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA
CAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG
GTCAGTTTTATTCCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
GAGAAACCACTTTTAGAACTATGAGTACTACTTTTGTAAATGTGAAAAACCCTCACAGAAAGTC
ATCGAGGCAAAAAGAGGCAGGAGTGGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTGACGTC
CACTGCTGGCTTTATTGAACAGCTAATAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTCATCCTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTTGTGCTGTATTCCATAATC
AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTATTTCATTCAC
AGAATGGAATTTTTTTGTTTCATGTCTCAGATTTATTTGTATTTCTTTTTTAACACTCTACATT
TCCCTTGTTTTTTAACTCATGCACATGTGCTCTTTGTACAGTTTTAAAAAGTGAATAAAATCTG
ACATGTCAATGTGGCTAGTTTTATTTTTCTTGTGTTTGCATTATGTGTATGGCCTGAAGTGTGGA
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAAATGTCACCAGACATTTGTATTATTT
TTATCATGAAATCATGTTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTATTTTGAATGC
ACAAAATGACTTAAACCATTATATCATGTTTCCTTTGCGTTCAGCCAATTTCAATTAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVL PQEAEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCGACCCGCCAGGAAAGACTGAGG
CCGCGGCTGCCCCGCGCGCTCCCTGCGCCGCGCGCTCCCGGGACAGAAGATGTGTCTCCAG
GGTCCCTCTGCTGCTGCGCTGCTCTGCTACTGGCCCTGGGGCTGGGGTGCAGGGCTGCCAT
CCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACCAGGTGCC
CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC
AGGCAGCTTTGCCGGCTGCCGGGCTGCAGCTCCTGGACCTGTCACAGAACCAGATCGCCAGCC
TGCCAGCGGGGTCTTCCAGCCACTCGCAACCTCAGCAACCTGGACCTGACGGCCAAACAGGCTG
CATGAAATCACCAATGAGACCTTCCGTGGCTGCGGCGCTCGAGCGCTCTACCTGGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCCCTCCTGGAGCTCAAGCTGC
AGGACAACGAGCTGCCGGCACTGCCCCGCTGCGCCTGCCCGCTGCTGCTGCTGGACCTCAGC
CACACAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACACTGCCAACGTGGAGCGCTGCGGCT
GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGCTCTTCAGCCGCTTGCACCACTCCACGACC
TGGATGTGTCCGACAACAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCTCCGGGGCTGACG
CGCTGCGGCTGGCCGGCAACACCCGCTTGGCCAGCTGCGGCGGAGGACCTGGCCGGCTGGC
TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCTTGCCTGGCGAGCTTCCGGCC
TCTTCCCCGCTGCGGCTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGC
TGGTTTGGCCCTTGGGTGCGCGAGAGCCACGTTCACACTGGCCAGCCCTGAGGAGACGCGCTGCCA
CTTCCCGCCCAAGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCAG
CCACCACCACAGCCACAGTGGCCACGAGGCGGCTGGTGGCGGAGCCACAGCCTTGTCT
TCTAGCTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCAGCCCGCCCTC
CACTGCCCCACGACTGTAGGGCTGTCCCCAGCCCCAGGACTGCCACCTGCCACCTGCCTCA
ATGGGGGCACATGCCACCTGGGGACACGGCACCACTGGCGTGTGTGTGCCCCGAAGGCTTCACG
GGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGCGCCAGCCCTACACCAGTCAAGCCGAG
GCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC
AGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG
GGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTGAGTACAGGTCACCCA
GCTGCGGCCCCAACGCCACTTACTCCGTCTGTGTCTATGCCTTTGGGGCCCGGGCGGGTGCCGAGG
GCGAGGAGGCTGCGGGGAGGGCCATACACCCCGAGCCGTCCACTCCAACACGCCCCAGTCAAC
CAGGCCCCGAGGGCAACCTGCCGTCTCTCATTTGCGCCCGCTGGCCGCGGTGCTCCTGGCCGC
GCTGGCTGCGGTGGGGGAGCCTACTGTGTGCGGCGGGGGCGGGCCATGGCAGCAGCGCTCAGG
ACAAAGGGCAGGTGGGGCAGGGCTGGGCCCTGGAAGTGGAGGGAGTGAAGGTCCCCTTGGAG
CCAGGCCCGAAGGCAACAGAGGGCGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCC
ACTCATGGGCTTCCAGGGCTGGCTCCAGTACCCCTCCACGCAAGCCCTACATCTAAGCCA
GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGTGCC
ACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCT
GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCC
CTAACGTCCCCAGAACCAGGTGCCATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC
CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCAC
TCCAGGCGGACCCTGGGGGCCAGTGAAGGAAGCTCCCGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAGAAACTGAAAGGAAGATGC
TTTAGGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCATTATTCT
GGGAAGATGTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTGTAAAGACAACGATGATATGAA
GGCCTTTTGTAAAGAAAAATAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVP LLLPL LLL LLL LALGPGVQGCPSGCQCSQPQT V FCTARQGT TVPRDVPPD TVGLYVFENGIT
MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLELKLQDNELRALPPLRLPRLLLLDLSSHNSLLALEPGILDANVE
ALRLAGLGLQLDEGLFSRLRNLDLDVSDNQLE RVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSLQALPGDLSGLFPRLRLAAARNPFNCVCP LSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALSSSLAPT WLSPTAPATEAP
SPPSTAPPTVGFVPQPQDCPPSTCLNGGTCHLGRHHLACLCEGFTGLYCESQMGQGT RPSPTP
VTPRPPRS LTLGIEPVSP TSLRVGLQRYLQGSSVQLRSLRLTYRNLSGPDKRLVTLRLPASLA EY
TVTQLRPNATYSVCVMPLGPGRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNIPLLIAPALAAV
LLAALAAVGAAYCVRRGRAMAAAAQDKGVGPGAGPLELEGVKVPLEPGPKATEGGGEALPSGSE
CEVPLMGFGPGGLQSPLHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCATGCG
GGTCCGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGGCCTCGGCGTCCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTCAGAAGAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGCACAG
AAGATATCAGCTTCTTAGAGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTA
CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCT
TTTCTTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAATGGAATGAAAATCCTTAATGGAAG
CAATAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCAAGGGACAGACTGCTCTTGG
CTTTCTGTATGCCTCTGGACTTGGTGTAAATCAAGTCAGGCAAAGGCTCTTGTATATTATACAT
TTGGAGCTCTTGGGGGCAATCTAATAGCCACATGGTTTTTGGTAAGTAGACTTTAGTGAAGGCT
AATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACTTTTTCAGCTTTTCATGATC
CAGATTTGCTTGTATTAAAGACCAAATATTCAGTTGAACTTCCTTCAAATCTTGTAAATGGATAT
AACACATGGAATCTACATGTAATGAAAGTTGGTGGAGTCCACAATTTTCTTTAAATGATTAG
TTTGGCTGATTGCCCCTAAAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTCTCTGAGTTG
GAATTGTCAGAATCATTTTTTACATTAGATTATCATAATTTTAAAAATTTTCTTTAGTTTTTCA
AAATTTTGTAATGGTGGCTATAGAAAAACAACATGAAATATTATACAATATTTTGCAACAATGC
CCTAAGAATTGTTAAATTCATGGAGTTATTTGTGCAGAATGACTCCAGAGAGCTCTACTTTCTG
TTTTTTACTTTTCATGATTGGCTGTCTTCCCATTTATTCTGGTCATTTATTGCTAGTGACACTGT
GCCTGCTTCCAGTAGTCTCATTTTCCCTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGG
AAGATTAATCATTTTTAATAAAATTATGTCTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA
AA

FIGURE 18

MRVRIGLTLLLCAVLLSLASASSDEEGSQDESLSKTTLTSDSVKDHTTAGRVVAGQIFLDSESEL
ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEPCHPFLFLDK
EYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEEAAKRRQMGEAEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERSYALLFGDYLPQNIQAAREMFEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTFGALGGNLIAMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEgyIEKSLDKLKGKNSYVNMD
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG
 CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGTAAAA
 GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA
 CATTGAGAAATATATTCCATGCTATCAGCTTTTGTAGCTTTTATAATTCTTCAGGCGAAGTAAATG
 AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATCCGT
 CGTCATTGAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACCTGCAGGAGCATTT
 TTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC
 ATCGACTGGAACATTCTTTATATAAACCTCAAAAAGGACTTTTTTCACAGGGTACCTTTTAGTGTT
 GCCAATCTGGGCATGTCTGAACAACCTGGGTTATAAACTGTATCAGGTTCCGTATGTCCACTGG
 TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTTGAAGAAGATGGATCCTTAAAGGAGG
 TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAGAGTATATGCAAAAAGTG
 GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAACGAGAAATTGA
 GAAAAGGAGAGGAGCACAGATTGAGGAGCAAGAGAGAAGAATCCAAAAGACCTCAGGAGA
 ACATTTTTCTTTGTGTCAGGCATTACGGACCTTTTTTCCAAATCTGAATTTCTTCATTCATGTGTT
 ATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAACACCAACCATCTCGATGTAGT
 AGACAATCTGACCTTAATGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCAC
 AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTA
 GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAAT
 GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC
 CTACATTTTGATCTTTTTAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAAC
 ATTTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTTACC
 TGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAACATCA
 GATGCTTTTATTTCCAAACCTTTTTTTCACCTTCTACTAAGTTGTTGAGGGGAAGGCTTACACAG
 ACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCAGCACT
 TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT
 GAGACCATGTCTATTAAAAAATAAAATGAAAAGCAAGAATAGCCTTATTTTCAAATATGGAAA
 GAAATTTATATGAAATTTATCTGAGTCATTAAATTTCTCCTTAAGTGATACTTTTTTAGAAGTA
 CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT
 AAAATTTAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLGGEVKGAEKNSITDSQMDDVEVVYTIQKYI
 PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFREERLLHKNLQEHFSNQDL
 VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
 QTHSSKFFEEGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA
 QIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHHLDVVDNLTL
 MVEHTDIPEASPASTPQIIKHKALDLDDRWFKRSRLLDTDKRSKANTGSSNQDKASKMSSPET
 DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCCAA
GCAGCGCGCAGCGAACGCCCCCGCCGCCACACCCCTCTGCGGTCCCCGCGGCGCCTGCCACCCCTTCCCTCCTTCCCC
GCGTCCCCGCTCGCCGGCCAGTCAGCTTGCCGGGTTGCTGCCCCGCGAAACCCGAGGTCACCAGCCCGGCGCTCT
GCTTCCCTGGGCGCGCGCCCTCCACGCCCTCCTTCTCCCTGGCCCGGCGCTGGCACCCGGGACCGTTGCCTGA
CGCGAGGCCAGCTCTACTTTTCGCCCGCGTCTCTCCGCTGCTCGCTCTTCCACAACTCCAACCTCTTCTCCC
TCCAGTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCGCTGCCGTAGCGCCGCTTCCCGTCCGGTCCCAA
GGTGGGAACGCGTCCGCCCCGGCCCGCACCATGGCACGGTTGCGCTTGCCCGCGCTTCTTGCAACCTGGCAGTGCTC
AGCGCCGCGTGTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGGCAGCTCTTACGTGTCCAAGGCTTC
AACAAGAACGATGCCCCCTCCAGAGATCAACGGTGATCATTGAAGATCTGCCCCAGGTTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAGTAAAGATGATTCAAAAGTGTGGTCAGCGAACAGTGCAATCATTG
CAAGCTGTCTTGCTTACGTTACAAGAAGTTGATGAATTCTTCAAGAAGTACTTGAAGATGAGAGAAATCCCTG
AATGATATGTTTGAAGACATATGGCCATTATACATGCAAAATCTGAGCTATTTAAAGATCTCTTCGTAGAGTTG
AAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATG
TTCCGCTGGTGAATCCAGTACCCTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAG
CCCTTCGGAGATGTCCCTCGCAAAATGAAGCTCCAGGTTACTCGTGCTTTGTAGCAGCCGCTACTTTCGCTCAAGGC
TTAGCGGTTGCGGAGATGTCTGTGAGCAAGGTCTCCGTGTAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCACTGCCGGGCTCTCGTACTGTGAAGCCATGTTACAAGTACTGCTCAACATCATGAGAGGC
TGTTTGGCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTATAGATGCTATGCTGATGTTGGCAGAGAGGTA
GAGGGTCCCTTTCAACATTGAATCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT
AATAGTGTTCAGTGTCTCAGAAGGTTTTCCAGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATTTCTCGT
TCCATCTCTGAAAGTGCCTTCAGTGTCTCGTTTACAGCCACATCACCCCGAGGAACGCCCAACCACAGCAGCTGGCACT
AGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATCTGGTCTCCCTTCCGAGCAAC
GTTTGAACGATGAGAGGATGGCTGCAGGAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC
CTGTTTGCAGTGACAGGAAATGGATTAGCCAACAGGGCAACAACCCAGAGTCCAGGTTGACACCAGCAAAACAGAC
ATACTGATCCTTCTGCAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGGAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG
TTTGACTACAATGCCACTGACCATGTGTTGGAAGAGTGCCAATGAGAAAGCCGACAGTGTGTTGTTCCCTGCGGCA
CAGGCCCTACCTCCTCACTGTCTTCTGCATCTTGTCTGGTTATGCAGAGAGAGTGGAGATTAATTCTCAAACCTGAG
AAAAAGTGTTCATCAAAAAGTTAAAGGACACAGTTATCATTCTTCTACCATCCTAGTGACTTTGCTTTTAAATGAA
TGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAAGTGTGACTTTGTTTCTCATTAGTTTGGG
AGGAAAAGGGACTGTGCATTGAGTTGGTTCTGCTCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGCATTGTGATTTTATCACTCTATTATTGTTTGTATGTTTTTTCTCATTTCGTTTGTGGGTT
TTTTTTTCCAACTGTGATCTCGCCTTGTCTTACAAGCAAAACAGGGTCCCTTCTTGGCAGTAACATGTACGTATT
CTGAAATATTAAATAGCTGTACAGAAGCAGGTTTATTATCATGTTATCTTATTAAGAAAAAGCCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH
LYMQNSELFKDLFVELKRYVVGNNLEMLNDFWARLLERMFLVNSQYHFTDEYLECVSKYTE
QLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRGL
VTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFQCGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNPEVQVDT
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
 CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT
 TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA
 TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCCGAGAGGCCACAGGGACCGAGGC
 CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCCTGAG
 AGCCCCGAGAAGAAAATTTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCTGTGATCATT
 TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA
 GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG
 AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
 TTCTTCTCCACCTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
 TGTTTTTCAAGATCATTTTGTGTTGCTCTCTCTAGTGTCTTCTCTCTCGTCAGTCTTAGCCT
 GTGCCCTCCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAACTGTAGCTTCCT
 AGCTAGTGTCAATTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT
 AAATGTCAAAAAAAAAAAAAAAAAA

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCACTCCTGCAGCTGCTGGTGTGCTTCTTACCCTGCC
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCGTGTCAAAAGCTACTTCCCCTACCTGA
TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG
ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
CACTTTTCACTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA
AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGTGCTG
CTCTGTGCAGAGCCCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
TCTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTC
GAGCCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA
GAACGCCCAGTTCTCCGAAATCCAAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG
GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCC
TTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTTA
GCAGAATGAGAGAAGACATTTCATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTC
CCAATGTTGTCCCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACAC
CCATGCGTCTCTAGGAAGTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
CCTCTCTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLLTLPPLHMLLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSSQIKGL
TGASGKVALLELGCCTGANFQFYPPGCRVTCCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFMWQQVFPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFSSKALICSFPSL
QLEQATHQPIYLPRLGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCCTATCCACCTCCCCAAGCCCCTTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
 GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCTCTTCTG
 CCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT
 CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTA
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAA
 GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTTCATTAAGAAATT
 GTTCCCTTCCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC
 AGTTAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 31

GTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
 CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
 ATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAATCATGTCGG
 GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGCTTTGGCCATGATGTTTACC
 TTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTCATTGGTTATTTTGGGATTGTTGTT
 TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
 CAGAAAGGGAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG
 CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATGACAGTTGAGCTTTTCCAAATCAC
 AAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCA
 TTTTCTTCTGGGTCCCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAAGTGCAGGAGCTGCCAGGTT
 ATGGAAGGCGGCCAAGTGAATATAAGCCCCTTTCGGGCATTTCGGTACATGTGGTTCGTACCATTT
 AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCCTGCCAGCAAATGACTATAGCTGGGGCAG
 TGGTTACTTGTATTTCACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTC
 TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTAAATCTCTGTGGTGAG
 GATTCCGAGAATCATGTGCATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT
 CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC
 AACCAGAATGCATATACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC
 ATTCAAAATCTTGTCCAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAA
 TTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC
 AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTTGCCTACTTAGTAGC
 CCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCCTGTGTTTGTCTGTGATC
 TGGAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA
 AAAAGGAGCAACAAATTAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
 GGGAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAAAACATT
 TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTT
 TTAAGACCTAATAAACCTATTCTTCCTCAAAA

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGGACCTCTCCCTGTTTCTTCTCTTAGA
ATAATTTGTATGGGATTTGTGATGCAGGAAAGCCTAAGGGAAAAAGAATATTCATTCTGTGTGGT
GAAAATTTTTGAAAAAAAATTTGCCTTCTTCAAACAAGGGTGTCTATTCTGATATTTATGAGGAC
TGTTGTCTCTACTATGAAGGCATCTGTTATTGAAATGTTCTTGTGTTTGTGCTGGTACTGGAGTAC
ATTCAAACAAAGAAACGGCAAAGAAGATTAAAAGGCCCAAGTTCACTGTGCCTCAGATCAACTGC
GATGTCAAAGCCGGAAAGATCATCGATCCTGAGTTTATTGTGAAATGTCCAGCAGGATGCCAAGA
CCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGTGGCGCTGCCG
TACACAGTGGTGTGCTTGATAATTGAGGAGGAAAAATACTTGTTCGGAAGGTTGCTGGACAGTCT
GGTTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACCACGATGGAGAGAATCCTT
TATCGTCTTAGAAAGTAAACCCAAAAAGGTTGTAACCTACCCATCAGCTCTTACATACTCATCAT
CGAAAAGTCCAGCTGCCAAGCAGGTGAGACCACAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAACTGCACAGCCGGTCACTCTGATGCAGCTTCTGGCTGTCACTGTAGCTGTGGCCACCCAC
CACCTTGCCAAGGCCATCCCTTCTGCTGCTTCTACCACCAGCATCCCCAGACCACAATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCCC
AGAGCTGATCCAGGTATCCAAAGCAAGATCCTTCAGGAGCTGCCTTCCAGAAACCTGTTGGAGC
GGATGTGAGCCTGGGACTTGTTCAAAAGAAGAATTGAGCACACAGTCTTTGGAGCCAGTATCCC
TGGGAGATCCAACTGCAAAATTGACTTGTCTGTTTTTAATTGATGGGAGCACCAGCATTTGGCAAA
CGGCGATTCCGAATCCAGAAGCAGCTCCTGGCTGATGTTGCCAAGCTCTTGACATTGGCCCTGC
CGGTCCACTGATGGGTGTTGTCCAGTATGGAGACAACCTGCTACTCACTTTAACCTCAAGACAC
ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTTCTAAT
GTAGTTCGGGCCATCTCCTTTGTGACCAAGAATTCTTTTCCAAAGCCAATGGAAACAGAAGCGG
GGCTCCCAATGTGGTGGTGGTGTGATGGTGGATGGCTGGCCACGACAAAGTGGAGGAGGCTTCAA
GACTTGCAGAGAGTCAGGAATCAACATTTTCTTCATCACCATTGAAGGTGCTGCTGAAAATGAG
AAGCAGTATGTGGTGGAGCCCAACTTTGCAACAAGGCCGTGTGCAGAACAAACGGCTTCTACTC
GCTCCACGTGCAGAGCTGGTTTGGCCTCCACAAGACCCTGCAGCCTCTGGTGAAGCGGGTCTGCG
ACACTGACCGCTGGCCTGCAGCAAGACCTGCTTGAACCTCGGCTGACATTGGCTTCGTCATCGAC
GGCTCCAGCAGTGTGGGGACGGGCAACTTCCGCACCGTCTCCAGTTTGTGACCAACCTCACCAA
AGAGTTTGAGATTTCCGACACGGACACGCGCATCGGGGCCGTGCAGTACACCTACGAACAGCGGC
TGGAGTTTGGGTTGACAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGC
TACTGGAGTGGTGGCACCAGCAGCGGGGCTGCCATCAACTTCGCCCTGGAGCAGCTCTTCAAGAA
GTCCAAGCCCAACAAGAGGAAGTTAATGATCCTCATCACCAGCGGAGGTCTTACGACGACGTCC
GGATCCCGACCATGGCTGCCCATCTGAAGGGAGTGATCACCTATGCGATAGGCGTGTGCTGGGCT
GCCCCAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGACCCTCCTTTTGTGGACGA
GTTTGACAACCTCCATCAGTATGTCCCCAGGATCATCCAGAACATTTGTACAGAGTTCAACTCAC
AGCCTCGGAACGTAATTGAGAGCAGGCAGAGCACCAGCAAGTGTGCTTTACTAAGTACGCTGTT
GGACCACCCACCGCTTAATGGGGCACGCACGGTGCATCAAGTCTTGGGCAGGGCATGGAGAAAC
AAATGTCTTGTTATTATTCTTTGCCATCATGCTTTTTCATATTCCAAAACCTTGGAGTTACAAAGA
TGATCACAACCGTATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTGTGGAGATTTTACAT
TTTGACAATTGTTTTCAAAATAAATGTTTCGGAATACAGTGCAGCCCTTACGACAGGCTTACGTAG
AGCTTTTGTGAGATTTTAAAGTTGTTATTTCTGATTGAACTCTGTAACCTCAGCAAGTTTTCAT
TTTTGTGATGACAATGTAGGAATTGCTGAATTAATGTTTAGAAGGATGAAAAATAAAAAA
AAA
AAA

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
 CQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
 ESFIVLESKPKKGVTPSALTYSSSKSPAAQAGETTKAYQRPPPIPGTTAQPVTLMQLLAVTVAVA
 TPTTLPRPSPAASTTIPRPQSVGHRSEQEMDLWSTATYTSSQNRPRADPGIQRQDPGAAAFQKP
 VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
 GPAGPLMGVVQYGDNPATHFNLKTHNTSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
 RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFEITIEGAAENEKQYVVEPNFANKAVCRTNG
 FYSLHVQSWFGLHKTQLQPLVKRVCDTDRDLACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN
 LTKEFEISDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGYWGGTSTGAAINFALEQL
 FKKS KPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAAWAAQEELEVIATHPARDHSFF
 VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
 431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTTTTCTGTTTTCAGAGTAGTTCACAACAGATCTGAGTGTTTTAATTAAAGCATGGAAT
ACAGAAAACAACAAAAAAGCTTAAGCTTTAATTTTCATCTGGAATTCACAGTTTTCTTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTTCCGCTGGCTGCTCTATCACGTGGTGTCTCCGACTACTCACCCGAGTGTA
AAGAACCTTCGGCTCGCGTGCTTCTGAGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCTTCCGAGTA
GGATGTCACTGAGATCCCTCAAATGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTTCTTTGTGATGTGGTAC
CTCAGCCTTCCCACTACAATGTGATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTTA
CAGACAAGACTTTCACCTCACACTTCGAGAGCATTCAAAGTGTCTCATCAAAATCCATTTCTGGTCATTC
TGGTGACCTCCCACCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAAGAAAAAGTCT
TGGTGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAGGAAGACAAAATGTTGGC
ATTGTCCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAATA
ACCTGACCTTGAAGACCATATGGCATTGAGGTGGGTAAGTGAAGTTTGCCCCAATGCCAAGTACGTAATG
AAGACAGACACTGATGTTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA
GAAGTTTTTACAGGTTATCCTCTAATTGATAATTTATCCTATAGAGGATTTTACCAAAAAACCCATATTT
CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTATATAATGTCCAGAGAT
TTGGTGCCAAGGATCTATGAAATGATGGGTACAGTAAAACCCATCAAGTTTGAAGATGTTTATGTCCGGAT
CTGTTTGAAATTTATTAAGTGAACATTCATATTCAGAAGACACAAATCTTTCTTTCTATATAGAATCC
ATTTGGATGTCTGTCAACTGAGACGTGTGATTCAGCCCATGGCTTTTCTTCAAGGAGATCATCACTTTT
TGGCAGGTCATGCTAAGGAACACCACATGCCATTATTAAGTTCACATTCACAAAAAGCCTAGAAGGACAG
GATACCTTGTGAAAGTGTTAAATAAAGTAGGTACTGTGGAATTCATGGGGAGGTCAGTGTCTGGCTT
ACACTGAACTGAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGG
CCCTTCAAGATGATATGTGGAGGAATTAATATAAAGGAATTGGAGGTTTTTGCTAAAGAAATTAATAGG
ACCAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAGGGTGTACTGAGTTATAAGCTCA
CTAGGCTGTAAAAACAAACAATGTAGAGTTTATTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTA
TATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAACTTCTTCACTGAAGTTATA
CTGAACAAAATTTTACCTGTTTTTGGTCATTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATT
ATTATTTAAATTACTTCACTTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAGGATAG
TGAATCATTCCTTACATGCAACATTTCCAGTACTTAACTGATCAGTTTATTATTGATACATCACTCCA
TTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTGGACTTTGTAAATATTTTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEA EKEDKMLA
LSLEDEHLLYGDIIRQDFLDYNNLT LKTIMAFRWVTEFCPNAKYVMKTD TDVFINTGNLVKYL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYFPKVFPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNIIHIPEDTNLFFLYRIHL DVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAGCCGCGGCAAGGATGGAGCTGGGTTGCTGGACGAGTTGGGGCTCACTTTTCTTCAGCTCCTTCTCATC
TCGTCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAATATCATGTGTCGGGAGTGCTGTG
AATATGATCAGATTGAGTGCGTCTGCCCCGGAAGAGGGAAGTCGTGGTTATACCATCCCTTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTTGAAAACGCAAGAGCTGCCGAAATGGCTCATGGGGGGGT
ACCTTGGATGACTTCTATGTGAAGGGTTCTACTGTGCAGAGTGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTCTGCGAGCCCCAAGGGTCAGATTTTGTGGAAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTTCATGC
TAAACCTGGGTTTGTATCCAACTAAGATTGTGCATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAG
GTTCTGTATGGAGACAACCGCATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG
GATCTCACTCCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTGACGGTTTCCATGCCATTATAGAGGAGATCACAGC
ATGCTCCTCATCCCTTGTTCATGACGGCAGTGCCTCTGACAAGGCTGGATCTTACAAGTGTGCTGCTTGGCAGGC
TATACTGGGCAGCGCTGTGAAAATCTCCTTGAAGAAAGAAATGCTCAGACCCTGGGGGCCAGTCAATGGGTACCAGAAAA
TAACAGGGGGCCCTGGGGTTATCAACGGACGCCATGCTAAAATGGCACCGTGGTGTCTTTCTTTTGAACAACTCCTATGT
TCTTAGTGGCAATGAGAAAAGAACTTGCCAGCAGAAATGGAGAGTGGTCAGGGAAACAGCCCATCTGCATAAAAGCCTGCCGA
GAACCAAAGATTTGAGACCTGGTGAGAAGGAGAGTTCTTCCGATGCAGGTTGAGTCAAGGGAGACACCATACACCAAGCTAT
ACTCAGCGGCCCTTCAGCAAGCAGAACTGCAGAGTGCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTACCCCTTACCAGCGCCCTGGGCAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGGGGGGACCATCTGTCATCCCTATCTGCGGGGAAAATGAGAACATCACTGCTCCAAAGA
CCCAAGGGTTGCGCTGGCGTGGCGTGGCAGGCAGCCATCTACAGGAGGACCAGCGGGGTGCATGACGGCAGCCACACAAGGGAGC
GTGGTTCTTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCCACTGTGTTACTGACCTGGGGAG
GTCACCATGATCAAGACAGCAGACCTGAAAAGTTGTTTGGGGAAATCTACCAGGATGATGACCGGGATGAGAAGACCATCC
AGAGCCTACAGATTTCTGCTATCATCTGCATCCCAACTATGACCCCATCTGCTTGATGCTGACATCGCCATCCTGAAAGCT
CCTAGACAAGGCCCGTATCAGCAGCCGAGTCCAGCCATCTGCTCGCTGCCAGTGGGATCTCAGCACTTCCCTCCAGGAG
TCCCACATCACTGTGGCTGGCTGGAATGTCTGGCAGAGTGAAGGAGCCCTGGCTTCAAGAACGACACACTGCCTCTGGGG
TGGTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGGAACCCACTGCCCTTCTGATATCTGCATGCAGAGACAGGAGGCATCGCGGCTGTGTCCTTCCCGGGA
CGAGCATCTCCTGAGCCACGCTGGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCCCTTACCAAGGTGCTGCCCTTTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCACTCCTTGAGAAG
TGTTTCTGTATATCCGTCTGTACGTGTTCATTGCGTGAAGCAGTGTGGCCCTGAAGTGTGATTGGCCCTGTAACCTGGCT
GTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGATGCCGCTCCA
CTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAAGAAGTAAGTTTCTTCAAGAAGACCATATACAAAACCTCTCCA
CTCCACTGACCTGGTGGTCTTCCCCAACTTTCAGTTATACGAATGCCATCAGCTTGACCAGGGAAGATCTGGGCTTCATGAG
GCCCCTTTTGAGGCTCTCAAGTTC TAGAGAGCTGCCTGTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCCTT
TGTGTACATGGCCACAGTACAGTCTGGTCTTTTCTTCCCCATCTCTGTACACATTTTAATAAAATAAGGGTTGGCTTCT
GAACTACAAAAA
AAAAA

FIGURE 39

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCCGTGATTTATTAACGTGGCTTAATC
TGAAGGTTCTCAGTCAAATCTTTGTGATCTACTGATTGTGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG
TTTGGGCCCTTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGCTTCTGTTC
TGGTCTTGCTTGGCTCAGTCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCTGTATTAGAACTCTGTA
AAGGTGCCTCCCACTACGGCCTGACCAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC
TCACAGCCACGGCTCCCTCCCCAGAGGTTCTGCAGTGCACCATCTCCTTAATGACAGACGAGCCTGGCCTAGACA
ACCTTGCCTACGTGTCTCGGCAGAGGACGGGAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGAACTA
GGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTCTTCGAA
GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACGAGGCGAGGAAAATCTGAAACACCACTGCCCTGAAG
TCTTTCCAAGGTTGTACACCTGATTCCAGATGGTGAATTAACAGCATCAAGATCAATCGAGTAGATCCCACTGAAG
GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCCATATCATTATCCAACACATTTATCGTGATGGGG
TGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
ACAACTACGCTGTGCTCTCTGCGGAGCCCTGCCAGGTGCTGTGGTGACTGTGATGCGTGAACAGAAGTTCCGCA
GCAGGAACAATGGACAGGCCCGGATGCCTACAGACCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC
CCGAGGAGCAGCTTGAATAAACTGGTGCAGAGGTGGATGAGCCTGGGGTTTTCATCTTCAATGTGCTGGATGGCG
GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTAGCCATCAATGGACATGATCTTCGATATGGCA
GCCCAGAAAGTGGGCTCATCTGATTGAGGCCAGTGAAGACGTGTTACCTCGTGTGTCGCCAGGTTCCGCGAGC
GGAGCCCTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCAGGGCCAGGGGAGAGGAGCAACA
CTCCCAAGCCCTCCATCCTACAATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAGACCCCGGTGAATCTCTCG
GCATGACCGTCGAGGGGAGCATCACATAGAGAATGGGATTGCTTATCTATGTCATCAGTGTTGAGCCCGGAGGAG
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTTTGTGAATGTGGATGGGGTCGAACTGACAGAGGTGAGCC
GGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACACATGGCCCCACCCAGTGACTGGTCCCCTACCT
GGGTGATGTGGCTGGAATTACACGGTGTGTATACTGTAAAGATATGTATTACGAAGAAACACAGCTGGAAGTC
TGGGCTTCTGCATTGTAGGAGTTATGAAGAATACAATGGAAACAAACCTTTTTCATCAAATCCATTGTTGAAGGAA
CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTGTGCTCAATGGTAGAAGTACATCAGGAATGA
TACATGCTTGTCTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTT
TTTTATAGAAATCAATGATGGGTGAGGAAAACAGAAAAATCACAAATAGGCTAAGAAGTTGAAACACTATATTTATC
TTGTGAGTTTTTATATTTAAAGAAAGAAATACATTGTAAAAATGTGAGGAAAAGTATGATCATCTAATGAAAGCCAGTT
ACACCTCAGAAAATATGATTCAAAAAAATTAAGTACTAGTTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTAC
AACATTGTTTATATTTTCTATTCAATAAAAAGCCCTAAACAACATAAATGATTGATTGTATACCCCACTGAATT
CAAGCTGATTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTAAATTTACAGCT
AAAAATTTTTTAAATGCATTGCTGAGAAACGTTGCTTTCATCAACAAGAATAAATATTTTCAGAAGTTAA

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTGCATCAAGTTCGCAATCAGATTGGAAAAAGCTCAACTTGAAGCTTT
 CTTCGCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCTTCAACCTGACT
 TTCCACCTTTCCCTACAAATTCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGTTGGGTGGGC
 CACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCTTAAAGCAAAGGAGTTCATGGCTAATTTCC
 ATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT
 GACAACCTGTCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACGAGATCTCAC
 TTTGGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAG
 CTTTACAGAGGGTCGCCATCCTCGTTCCCCACCGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
 CATCTGCATCCCTTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTATCCACCAGGCTGAAGG
 TAAAAAGTTTAATCGAGCCAAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG
 ACTGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG
 CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTTGG
 GGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACTACTGGGGAT
 GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTG
 CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
 GATGAAGCTCTTACACCAAGTGTACAGAGTCTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAAT
 TAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGGATTCTCGTTTGGTGCATGACCC
 TGGATCTTTTGGTGATGTTTGAAGAAGTATTCTTTGTTTGAATAATTTGGCCTAGAGACTTCAA
 ATAGTAGCACACATTAAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTCCTTTTTGATTTTCT
 TAGCAGAGCTCCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGAT
 CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
 AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGAT
 AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA
 AGGTACGAAGATACAATACTGTTATTCAATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTCAAGT
 GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTTGGGAATGAAGA
 GGTAGCAGGAGGGTGGAGTGTGGCTGCAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
 CTTCAGGGGAGGACCTGCCCAGGTATGCCCTCCAGTGATGCCACCAGAGAATACATTCTCTATTAGT
 TTTTAAAGAGTTTTTGTAAAATGATTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT
 ATTAACTAATAATAAATATGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAAGCAAAA

FIGURE 42

MGFNLTFLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVGRNSTGYRLRYSGYFGGVLTALSREQFFKVNGFSNNYWGWWGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG
 CCGCTTGCCCTCCTGCTCCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCACAACAG
 ACGGGACAACCTTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT
 GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCTGCGGCTGCTGTC
 ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCGTCCTTCCC
 TTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGAATAAAATGGCTGGTTCTTTGTTT
 TCCAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI
CIFCCGCCHRSKCGMCKKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
 TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCLISIQALLKGPLMCNPSNSNANCEFSL
 KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
 LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA
 AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
 ATCCGTGGGCTGCAGACCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGA
CATGGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAGCCAATGACC
 CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
 TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
 CAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCCTCATCACTCCAGGCTCTGCCA
 CTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
 TCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCCCT
 TTCTGATCAGGAGGCTTCTTTATGAATTAACTCGCCCCACCACCCCTCA

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
 AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
 AGGAGACCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG
 GGGCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGGAGGCTGGCCCCCTGCAGAG
 CGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGACATGGCCTGGGAGACGCCCTGA
 GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGGCAGCTGGCTCTAAAGTCAGTGAG
 GCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC
 AGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGA
 TTGGCAGACAGGCAGAAGATGTCATTGACACGGAGCAGATGCTGTCCGCGCTCCTGGCAGGGG
 GTGCCTGGCCACAGTGGTGCTTGGGAAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG
 CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCG
 GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGA
 GGGCCACCAAACCTTTGGGACCAACACTCAGGGAGCTGTGGCCCAGCCTGGCTATGGTTCACTGAG
 AGCCAGCAACCAGAATGAAGGGTGACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
 ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC
 AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG
 TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT
 CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGA
 CATAAACCCGGGTGTGAAAAGCCAGGGAATGAAGCCCGGGGAGCGGGGAATCTGGGATTCAGGG
 CTTAGAGGACAGGGAGTTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATCGCCTCCTTG
 GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT
 GGTGGAGTCAATACTGTGAACCTCTGAGACGTCTCCTGGGATGTTAACTTTGACACTTTCTGGAA
 GAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGAAGCTCTC
 GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCACACTCCCTCCTTAA
 AACACCACCTCTCATCACTAATCTCAGCCCTTGCCCTTGAAATAAACCTTAGCTGCCCCACAAA
 AA
 AA

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGCTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAAGTCCGCGCGCTCCAGTGTTTCC
CACAGCCCCCAAACGGAAGTGGTTTTGGGGTCACCTGGGCCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGTATGGCTGGGTCCCAT
CATCCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCAGTCCCTCAGCTGCCA
TTGCACCCAAGGATAATCTCTTCATCAGGTTCTTGAAGCCTGGCTGGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCCGATGCTGACGCGCGCTTCCATTTCAACATCCT
GAAGTCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGAGGCTCAGTGCCCTTGTAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT
ATTACCTCTCCCATGACGGGCGCGCTTCCACAGGGCCTGCCGCTGGTGCATGACTTCACAGAC
GCTGTCATCCGGGAGCGGCGTTCGACCCCTCCCACTCAGGGTATTGATGATTTTTTCAAAGACAA
AGCCAAGTCCAAGACTTTGGATTTCAATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCAG
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCC
AGCTGCCCTTCTTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGACC
CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTTATTCTTTCTCCGCA
GGGCCCAGGAAGTGCATCGGGCAGGCGTTCCGCATGGCGGAGATGAAAGTGGTCTGGCCTTGAT
GCTGCTGCACTTCCGGTTCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCAGTGAATTTCTGAC
CCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGTGTCAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPKRNFWGHLG
 LITPTEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
 WLGEILLSSGGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
 SLMTLDSLQKCIFFSDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
 RLVDFTDAVIRERRRTLPTQGIDDFKDKAKSKTLDLFDVLLSKDEDGKALSDEDIRAEADTF
 MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
 LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP
 LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
 VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA
 GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTTCTGCCT
 TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCATTGACGTGG
 TACAGCCTTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTTGCCGTGTGTCTTGC
 ATAATTTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACTATGGACAGAAGCTGGTGGACAGTTTT
 GTAACATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGCAGCAATGTGTTGCTT
 GTGATTGGAACATTTGAGGGTTACTTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
 GCACAGGATGAGAAGTGGGTCTGTATCTTGTGGAGTGAATCTTCCTCATGTACCTGTTTCCTC
 TCTGGATGTTGTCCCACTGAATCCCATGAATACAAACCTATTTCAGCAACAGCAAAAAAAAAAAAA
 AAA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

[illegible]

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSI FKLSVFIPSQEFSTYRQWKQKIVQAGDKDLDG
 QLDFEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMDKNG
 TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTPDEFTVEERQTGMWWRHLVAGGG
 AGAVSRTCTAPLDRKVLMMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIAK
 FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
 ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNALQHYAVNSADPGVFVLLACGTMSSTC
 GQLASYPLALVRTRMQAASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
 VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCCATGGC
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTTGGTATTTAGGGAGACACTCCATCACAGTCACTACTGTCGCCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAAACCTTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTGCGAGCAGGATGAAATGTTTACAGAGCCGGACAGCAGTGTTTGCTGATCAAGTGATAGTT
GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCTCC
CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACTTCTCGGAAGTCTCCAATAC
CAGCTTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG
ACAGAATCGGAGATCAAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGT
CTCTTCTTTCTTTGCCATCAGCTGGGCACCTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAAAT
GTGCCTTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAATATTTACAC
CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAAGCAGAAGGCTCCAATATGAACAAGATAAAT
CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTCATGTGAAGTAGACAAGTGTGTTAAGA
GTGATAAGTAAATGCACGTGGAGACAAGTGCACTCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGGAGTGAGAGGACAGGATAGTGATGTTCTTTGTCTCTGAATTTTITAGTTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATCCAC
AAATTAAGCTGTAGTATGTACCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG
GCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCTTGCTTTC
TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT
CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTAAACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
 DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
 YIITSKKGKNANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
 NTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTSEIKRRSHLQLLNSKASL
 CVSSFFAISWALLPLSPYLMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
 220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCCTGGTGGTGGTTATGGGGGTCCCTGCCCTGGAGGGCCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCTCTGGAACCTCAGGAGGACCATATG
GCGGTGCAGCTCCCGGGGGCCCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGCCCAGCAGCCT
GGGCTTTATGGACAGGGTGGCGCCCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTTCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGGGAAGAACCCTCTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
AAATGGGCTACAACCTGAGCCCCAGTTACCCAGCTTCTGGTCTCCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA
CCATGACAGCTTCTCGGATGCTATGACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCCTGCATCATAGCCACCA
AATAGTGAGGACCGGGGCTGAGGCCACACAGATAGGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTTGTAAATGG
AGTTAGTGTCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGAGTGTTCATCGGCCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCCAAAGTGGAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCTGCCACACCCATAAAT
CCTTGTGTGTTAACTTCTAGCTGCCTG⁶GGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT
CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCCCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTTAATTTGCATTTTTTTTC
ATTTGGGGCCAAAAGTCCAGTGAAATGTAAAGCTTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
 HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVDSH
 SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR
 DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
 DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
 66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
 TCCCAGGCCTCTTTGCCCGGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCCAAACCTTCGGGACC
 AACTTGCCCTCAGCTCGGACAACCTTCCTCCACTGGCCCCCTCTAACTCTGAACATCCGCAGCCCGC
 TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCCTCTGAAGCTCAGCGTGCCTCCATCAGATG
 GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCCTGCCATG
 GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCTGGGGGA
 AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGCTGCGGCCCTCGCTCCGGGCAGTGGCCCTT
 TGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
 TCGGAGTCCAGACGACTGCCCCGTCTAATTCACTGGGAGCCGGGGGAAAAATCCTTTCCCAACG
 CCCTCCCTGGTCTCTCATCCACAGGGTTCGCTGATCACCCCTGGGGTACCCTGAATCCAGTG
 TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACACCCTGAGGGA
 ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG
 CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTAATCGGTATCCAGGAG
 GCAGCTGGGGGAATATTCATCTATACCCAGGTATCAATAACCCATTTCTCCTGGAGTTCTCCGC
 CCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTTCCTAATCCTCCAAGCCCTAGGTTGCAGTG
 GGGCTAGAGCACGATAGAGGGAACCCCAACATTGGGAGTTAGAGTCTGCTCCCGCCCCCTTGCTG
 TGTGGGCTCAATCCAGGCCCTGTTAACATGTTTCCAGCACTATCCCCACTTTTCAGTGCCTCCCC
 TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AA

FIGURE 64

MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGSPNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSQVQRWPPSWGLPAMDSWPPEDPWQMAAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNISLGAGGKILSRP
PWSLIHRVLPDHPWGTNLNPSVSWGGGGPGTGWGTRMPHPGEGIWGINNQPPGTSWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPFPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTTATGGGGTC
 TGGGCTGCCCCCTTGTCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA
 CTTTGCAACTGAAGCTGAAGGAGTCTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA
 TTGCTTGAAAAGCTCTGCCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTCCACCATGC
 AAGATCTCAACACCATGTTGTCTGCAACACATTGACAGCCATTGAAGCCTGTGTCCTTCTTGGCCC
 GGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAGGCCCCCACCTC
 CTGAGTGGCAATAAATAAAATTCGGTATGCTG

FIGURE 66

MGSGPLVLVLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 68

MANPGLGLLLLALGLPFLRLARWGRAWGQIQTTSANENSTVLPSSSTSSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLRKLRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTGAGAGGTTTTGTTTTCTCTTAGTTCTGTGCCTGCTGCACCAG
 TCAAACTACTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATATTGTCATGTTATAGATCCTAGTGTGCCAGAA
 GATGAAAAATAAATTGAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGA
 TTTTTTTTCAAAAATGTATCTATATTAATTCCTGAGAATTGGAAGGAAAAATCCTCAGTACAAAAGGCCAAAAATGAA
 AACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCATACACCAAGCAGTTTACA
 GAATGTGGAGAGAAAGGCGAATACATTCACCTTCAACCTGACCTTCTACTTGGAAAAAACAATAATGAATATGGACCA
 CCAGGCAAACTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTGATGAGTACAATGAAGATCAGCCTTTC
 TACCGTGCTAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAAGTGT
 CAAGGAGGCAGCTGTCTTAGTAGAGCATGCGAATTGATTCTACAACAAACTGTATGGAAAGATTGTCAATTCCTT
 CCTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATGCAAAGTATTGATTCTGTTGTTGAATTTTGTAAACGAA
 AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAGTGCAATTTTGAAGTACATGGGAGGTGATTAGCAAT
 TCTGAGGATTTTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATCAGTCAA
 AGAATGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGTAAGGACCGCTAAATCGAATGAATCAAGCAGCA
 AAACATTTCTGCTGCAGACTGTTGAAAATGGATCCTGGGTGGGGATGGTTCACCTTGTAGTACTGCCACTATTGTA
 AATAAGCTAATCCAATAAAAAGCAGTGATGAAAGAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA
 ACTTCCATCTGCTCTGSAATTAATATGCATTTAGGTGATTGGAGAGCTACATCCCAACTCGATGGATCCGAAGTA
 CTGCTGCTGACTGATGGGGAGGATAACACTGCAAGTTCTTGATTGATGAAGTGAACAAAGTGGGGCCATTGTTTCAT
 TTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTTTATGTT
 TCAGATGAAGCTCAGAACAATGGCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAAAG
 TCCCTTCAGCTCGAAAGTAAGGGATTAACTGAATAGTAATGCCTGGATGAACGACACTGTCAATATGATAGTACA
 GTGGGAAGGACAGTCTTTCTCATCATATGGAACAGTCTGCCTCCAGTATTTCTCTCTGGGATCCAGTGGAAACA
 ATAATGGAAATTTTACAGTGGATGCAACTTCCAAATGGCCTATCTCAGTATTCAGGAACTGCAAAAGTGGGCACT
 TGGGCATACAATCTTCAAGCCAAAGCGAACCCAGAAACATTAACTATTACAGTAACCTCTCGAGCAGCAAAATCTTCT
 GTGCCTCCAATCAGAGTGAATGCTAAATGAATAAGGACGTAACAGTTTCCCGAGCCCAATGATTGTTTACGCAGAA
 ATTTACAAGGATATGTACCTGTTCTTGAGCCCAATGTGACTGCTTTCATTGAATCAGAGAATGGACATACAGAAGTT
 TTGGAACCTTTGGATAATGGTGACGGCTGATTCTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATAT
 ACAGAAAATGGCAGATATAGCTTAAAGTTTGGGGCTCATGGAGGAGCAACACTGCCAGGCTAAAATTACGGCCTCCA
 CTGAATAGAGCCGCGTACATACCAGGCTGGGTAGTGAACGGGGAAATGGAAGCAAAACCGCCAGACCTGAAATTTGAT
 GAGGATACTCAGACCACCTTGGAGGATTTAGCCGGAACAGCATCCGGAGGTGCATTTGTTGATCACAAGTCCCAAGC
 CTTCCTTGCCTGACCAATACCCACCAAGTCAATCAGACCTTGATGCCACAGTTCATGAGGATAAGATTATTCTT
 ACATGGACAGCACCAGGAGATAATTTGATGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTGAAGTATTCTT
 GATCTAAGAGACAGTTTGTATGCTCTTCAAGTAAATACTACTGATCTGTCAACAAAGGAGGCCAACTCCAGGAA
 AGCTTTGCATTTAAACCGAAAAATATCTCAGAAGAAAAATGCAACCCACATATTTATTGCCATTAAAGTATAGATAAA
 AGCAATTTGACATCAAAAGTATCCAACATTGCACAAGTAACCTTTGTTTATCCCTCAAGCAAACTCTGATGACATTGAT
 CCTACACCTACTCCTACTCCTACTCCTGATAAAAGTCATAATCTGGAGTTAATATTTTACGCTGGTATTG
 TCTGTGATTGGGTCTGTTGTAATGTTAACTTTATTTTAAGTACCACCATTTGAACCTTAACGAAGAAAAAATCTTC
 AAGTAGACCTAGAAGAGAGTTTTAAAAAACAAAACAAATGTAAGTAAAGGATATTTCTGAATCTTAAAAATTCATCCCAT
 GTGTGATCATAACTCATAAAAATAATTTAAGATGTCGGAAGGATATTTGATTAAATAAAAACACTCATGGATA
 TGTAAAACTGTCAAGATTAAATTTAATAGTTTCAATTATTGTTATTATTATTGTAAGAAATAGTGATGAACAAAG
 ATCCTTTTTCATCTGATACCTGGTTGTATATTATTTGATGCAACAGTTTTCTGAAATGATATTTCAAATTTGCATCAA
 GAAATTTAAATCATCTATCTGAGTAGTCAAAATACAAGTAAAGGAGAGCAAAATAAACAACATTTGGAAAAA
 AA

FIGURE 70

MGLFRGFVFLVLCLLHQSNTSFIKLNNGFEDIVIVIDPSVPEDEKIIIEQIEDMVTASTYLFE
 ATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
 IHFTPDLLLGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN
 RVYKCGGGSCLSRACRIDSTTKLYGKDCQFFPDQVTEKASIMEMQSIDSVVEFCNEKTHNQEAP
 SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPVFSLLKISQRIVCLVLDKSGSMGGKDRNLNR
 MNQAAKHFLLOTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTICSIGIK
 YAFQVIGELHSQLDGSEVLLLTGDNATSSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
 SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL
 ITWNSLPFSLWDPSTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
 AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
 GADSEKNDGVYSRYFTAYTENGRYSLKVRAGGANTARLKLRLPLNRAAYIPGWVVNGEIEANPP
 RPEIDEDTQTTLDFSRASGGAFVVSQVPSLPLPDQYPPSQITDLATVHEDKIIILTWTPAGDN
 FDVGKVQRYIIRISASILDRLDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI
 KSIDKSNLTSKVSNIQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVI
 VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

TCCTCTTAGGTGGAAACCCCTGGGAGTAGAGTACTGACAGCAAAAGACCGGGAAAGACCATACGTCCCCGGGCGAGGGGTGA
CAACAGGATGTCATCTTTTGGATCTCGTGTGTGGCTGCCTTCTCTATTTCAGGAAAGAACGCCAAGGTAAATTTTGACCCA
GAGGAGCAATGATGTAGCCACCTCTTAACCTTCCTTTTGAACCCCGAGTTATGCCAGGATTTACTAGAGATGTCA
ACTCAACGACGACGCGGCTCGGTTTCGGCTTAACTTGTGTGTGGAGGAGAGACCTTTGTGGGGCTGCGTTCCTTTAGCA
GTGCTCAGAAGTGACTTGCCTGAGGTGGACGAGAAGAAAGGTCCTCTTGTGCTTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTGAAAATCTGGAGATTTCACTTCAGTCATTTGCTTCGCTGCCGCAAGATCATCTTTTAAA
AGTAGAGAAAGCTGCTCTGTGTGGTGGTTAACTCAAGAGGCGAGAACTTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCCAACGCGATGCTTCTGTGGTTAGCCAGGAAAGCCCTTCGCTGGGGGCCCGGCTTTTGGAGGATGCC
ACCGGTTTCGAGCAGTGGCTGATTCTCTGAAATGATGATGGTTGCGCGGGGGCTGCTTGCCTGGATTTCCTGGGTGGT
GTTTTGCTGGTGCCTCTCTGCTGTGCTATCTCTGTCTGTACATGTGGCTGCACCCCAAAGGTCGACGAGGAGCAG
CTGGCATCTCCGAGGCCCCAACAGCCCCAGGGGAAGGAGGTTACAGGCGCTCTCTCAGGATGGGAGGAGCAGCAC
CGCAACTACGTGAGCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGG
AATGGGCGATCAACCAAGCCAGCATGTGCTGGCTGGGTCTGGACAGGAGGCCCGCCAGAGAAAACAGCGCCAGCTC
CTGGCCTTCTCGACTCGAGTGGTGGACAGGACAGGTGAATGCTGGCGCTCAAGCTGGCCACAGGATGACGAGATG
CCTTTTCGATAGCTTTTACTTCTACAGAAGTGTACACAGCTGGAGACTGGCCTTACC CGCCACCCCGAGGAGAAGCCTGTG
AGGAAGGACAGCGGGATGAGTTGGTGGAAAGCCATTGAATCAGCCTTGGAGACCTTGAACATCTCGACAGAACAGC
CCCAATCACCGTCTTACACGGCCTTGATTTCATAGAAGGATCTACCGAACAGAAAGGGACAAAGGGACATTTGAT
GAGCTCACCTTTCAAAGGGGACCAACAAACAGAAATCAAACGGCTCATCTTATTTGCACAGCTCAGCGGCTCATGAA
GTGAAAAATGAAAAGCTCAACATGGTGGCCAACACGCTTATCAATGTTATCTGCTCTAGCAAAAAGGCTGGACAATGTA
CGGCAGTTCATCGAGAAATTCAGGGAGATGTGCTTATGACAGGATGGGAGAGTGCATCTCACTGTTGTTTACTTTGGG
AAAGAAGAAATAAATGAATGACGAAGAATACTTGAACAACTTCAAAGCTGCCAACTTCAGGAAGCTTTACCTTCTATC
CAGCTCAATGGAGAAATTTCTCGGGAAGAAGGACTGATGTTGGAGCCGCTTCTGGAAGGGAAGCAACGCTCTTCTC
TTTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAG
GTATTTTCTCCAGTCTCTTTTCAGTCAGTACAATCTGGCATATAATACGCCCCACATGATCGATGCTCCCTTCTTGGAA
ACGACGCTGTGCTATAAGAAAGGAACTGAAATTTGGAGAGACTTTGGATTTGGGATGACGTGTGATGTCGTCAGCAG
TTCATCAATATAGTGGGTTTGATCTGGACATCGAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTC
CAGACCAACCTCATAGTGGTACGAGGCGCTGTGCGAGGACTTCCACCTCTGGCATGGAGAGGCTGTCATGGGACGAG
TGACCCCCGAGCAGTACAAGATGTGCATCGACTCAAGGCCATTAACGAGGCACTCCACGCCGAGCTGGGCTGCTG
GTGTTCCAGGCAACGAGATAGAGGCTCACCTTCGCAACAGAAACAGAAAGCAAGTAGCAAAAACATGAATCACTCCCA
GAAGGATGTGGGAGACACTTTTTCTTCTTTTGAATTAATCTGAAATGGCTGCACAGAGAAAAGACTTCCATAAA
GGACGCAAAAAGATTTGACTGATGGGTGAGAGATGAGAAAGCCCGATTTCTCTGTGTGGGCTTTTACAACAGA
AATCAAAATCTCCGCTTTGCTTCGAAAGTAAACCGATGTGCACCTCTGGAAGTGCTGCAAAAAGGCAAGATGCTTTG
AGATTTATAAGCCTAATGTTGTGAGGCTTTGATGGTGTTTACAATACTGAGACTGTGTTTGTGTGCTCATTTGA
AATATTTCATGATTTAAGAGCAGTTTTGTAAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCTCATATGAATGA
GCCATTCAGCAGGCGCTTAGTTTCTAGGAATGCTAAAAATCAGAAGGCGAGGAGAGGATAGGCTTATTATGATACT
AGTGAGTACATAAGTAAAAATAAATGGACCAAGAAAAGAAAGCAACATAAATATCGTGTATATTTTCCCCAGAT
TAACCAAAAATAATCTGCTTATCTTTTGGTGTCTTTTAACTGTCTCCGTTTTTTTCTTTTATTTAAAAATGCACT
TTTTTTCTTTTGTGAGTTATAGTCTGCTTATTTAATACCATTTTGCAAGCCTTACAGAGAGACCAAGTTGGCCTAC
ATTTTTATATTTTAAAGATACTTTGAGATGCATTTAGAAGTTTCAGTTCAGTAACGATCAAAATGATGCCATAT
CCAAGGACGTCCTCAATGCTGATTCTGTGAGGCACTGAATGTGAGGATGTGAGACATAGGGAAGGAATGGTTGTGAT
AATACAGACGTACAGATACTTTCTCTGAAGATTTTTCGAAGGAGGAGCACTGAACACTGGAGGAAAAGAAATGAC
ACTTCTGCTTTCAGAAAAGGAAACTCATTCAGACTGGTGATCTGTGATGTAACCTAAAAGTCAGAAACCAATTTT
CTCTCAGAAGTAGGGACCGCTTTCTTACCTGTTTAAATAAACCAAGATATCCGTTGTGAACCAACAACTCTTTTCT
AAAACAGGGTGCTCCTCTGCTTCTGGCTTCGCTACATAAGAAAGATGAGAAAAATATATATATATATATATTGT
GAAAGATCAATCTCTGCGCAAGATCTAGTGGGATGGAAGTTTTTGTACATGTTATCCACCCAGGCCAGGTGGAAG
TAACATGAATTTTAAAAATTAAGCACTTCTACTCAATCAACGAAGTGCTTGAATAATGCAATTTTATTTACCAATTT
CAACTTATTTTAAAAATAAATCAGTTAACTAGAGTGTTTTCTTCAATCATGTGAAATTTATGCGCAGCACAG
ATGCATGAGCTAATTATCTCTTGTAGTCTTGTCTCTGTTTGTCTACAGTAAACTCATTGTTTAAAGCTTCAAGAAC
ATTCAGCTGTCTGGTGTTGTTAAAAATGCATTTGATTTGATTTGATCTGGTAGTTTATGAAATTTAATTTAAACACAGG
CCATGATGGAAGGTGGTATTTGCACAGCTATAATAAATGATTTCGTGATGAA

FIGURE 72

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETNNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLIELTFKGDHKHEFKRLILFRPFSPIMKVNEKLNMAN
TLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLF
SQYNPGIYGHHDVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDLIDIKGWGGEDVHLYR
KYLHSNLIVVRTPVRGLFHLWHEKRCMDELTPQYKCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHPASYPECGNNAQSPIDIQTDSVTFDPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGTGTC
 TCTGGTGGTTTGCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACA
 ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA
 ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCTGACAGCTCC
 AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATACTCCAATCTGA
 AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
 ACGCTGGTGTCTACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCTTCGTCCC
 AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT
 CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTCCTTTT
 TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT
 GATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAAATCGTGATTA
 ACTTTATCACCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGA
 AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
 GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAG
 AAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAATACCCCCGGATAAA
 ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTTGTGCGGGGCTGAAGAGCAGGA
 GCTCAGTTTGAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG
 TCTTGGGCCCCGAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCCCTGGCG
 CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGA
 TCCCCAAACTGGCAGGCTGTGTATTCCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCG
 AGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT
 CCAGACAGGCCACCAGGAGAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGTTATA
 TGTGCAGATGGAAAACTCATGCCAACACTTCCTTTTGCCCTTTTGTTCCTGTGCAACAAGTGAG
 TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTCTAGTGT
 CTGTGAGAATTACTTATTTCTTTCTCTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA
 GGTCTCTTAACAATGATGGTGGGCCCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA
 AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTCAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSSISQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
 SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLK
 DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNFDRFFVPAEK
 IVINFITLNISSDKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
 DSEENTEGTSLTQQESLSRTIPDKTVIEYEYDVRTTDCAGPEEQELSLQEEVSTQGTLLSQ
 ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
 EGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDNATSILQQPLLSAMREK
 PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKATIVE
 FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
 VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLKLLDSQGKVT
 KWFNNSAASLTMPITLDNIPFSLIVSQDVVKAABAVALSPEEFMVLLDSVLPESAHLKSSIGLIN
 EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTLGIEASSEAQFYT
 KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSVGVPSLVKALG
 FEAAESSLTkdalvltpaslwkpsspvsq

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
 TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT
 CTTGGCCTCCAACTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT
 GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTGGTGCCAGCATTGTGACAGCAGTTGGCTTCT
 CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
 ACCCTTCTGGGCCTGCCCCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAAT
 CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
 GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTCATCCTTGGAGGCCTCTCGGGATTCT
 ATTCCTGTTGCCCTGGAATCTTCATGGGATCTACGGGACTTCTACTCACCAGTGGTGCCCTGACAG
 CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG
 CTGGAATCATCCTCTGCTTTTCTGCTCATCCCAGAGAAATCGCTCCAATACTACTACGATGCCTAC
 CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA
 GTTCAATTCCTACAGCCTGACAGGGTATGTGTGAAGAACCAGGGGCCAGAGCTGGGGGGTGGCTG
 GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT
 GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
 GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC
 TCACCTTGCTGCTCCCCTGCCCTAAGTCCCCAACCTCAACTTGAAACCCATTCCCTTAAGCCA
 GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA
 TCCCACTGACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT
 GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC
 CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCACTCTTGTTATGACTCCACAGTGTCCA
 GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
 CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGLQLVGYYLGLLGLLGLTLVAMLLPSWKTSSYVGASIVTAVGFSGKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMTSSAIISSLACIIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAVNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPD MFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPVHCPQPVTEPQCCPKC VEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPPTPAPTGLSAPLSFI PRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTY S
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQ RVTCPT EYPCRHPEKVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCACGCAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTCTAATCCATCCGTACCTCTCTGTCA
TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGTTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGAG
GACGCAGCATTCTCCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATC
AAGGCAGGACAAAATCGGTGAAGGATTCTATTGCGGAGGGGGCGCATCTCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCCACTCTTACTACCAGAAGGCCATCTGGGAGCT
ACAGGTGTCAGCACTGGGCTCAGTTCTCTCATTTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCACTCCTCGGGCTGGTTCCCCCGGGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTGTCC
ACAGACTCCAGGACAAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAA
CGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACCTTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGCCCTA
TTTTTTGGCATTGTTGGACTGAAGATTTTCTTCTCCAAATCCAGTGGAATAATCCAGGCGGAACTGGACTG
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGAAACACGCAGTGGAGGTGACTCTGGATCCAG
AGAGGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTGAACCATAGAAAAGCTCCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCGTGGGAGTGTGCCGGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTC
ACATTAATCCCCGTTTTATCAGCGTCTTCCCCAGGACCCACCTACAAAAATAGGGGTCTTCTTGACTA
TGAGTGTGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTCGGTTTG
AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAAATGGAACCTCCCATAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGCAATCCCAGAGACAAGCAACAG
TGAGTCTCTCTCACAGGCAACCACGCCCTTCTTCCCCAGGGGTGAAATGTAGGATGAATCACATCCCACAT
TCTTCTTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCCA
GATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGTGTCTATGGCTGCCCTGAGCTGGGAGGGAAGAAGG
CTGACATTACATTTAGTTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAG
AACCCTCAGGAATTCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCCTAATGGTTTGTTCATTATATTACACTTTCAGTAAAAAAA

FIGURE 84

MALMLSLVLSLLKLGSQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSVVH
LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAglyGCRISQSYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFPPTAKWKGPQGDLSTDSRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAEldWRRKHGQAEldARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQSfQAGKHYEVDGGHNKRVRVGVCRDDVDRRKEYVTLSPDHGyWVLRlNGEHLYFT
LNPRFISVFPRTPPTKIGVFLDYECGTISFFNINDQSLIYTLTCRfEGLLRPYIEYPSYNEQNGT
PIVICPVTQeSEKEASWQRASAIpETSNSESSSQATTpFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCTGCCCCT
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCTGTA
CGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCTACCCCTCGCATGGCTGGATT
TACCCTGGCCCAGTAGTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGGACCCACATACCAAGAATTGCACCCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA
TACTTCTTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCCAGGCACCCCTGGAGTCCGGCTGCCCCC
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCTCGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCGCTCCTCGGTGCTCACCTCATCCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTGAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCCTACCCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGA
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCT
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCTGTGCCCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCTCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTTCGGGGGAGCTGGAG
CCACAGCCCTGGTCTTCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTC
AGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCCAG
CTTCTGCCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC
TCCCTTTTATTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLPLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIYPGPVVHGYWFREG
 ANTQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY
 KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVSPDPSTTRS
 SVLTLPQPDHGTSLTCQVTFPGASVTTNKTVHLNVSYPQNLMTVFQGDGTSTVLGNGSSL
 SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
 LGSQQVYLNVSLSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDGTGIE
 DANAVRGASQGPLTEPWAEDSPPDQPPASARSSVGEGELQYASLSFQMKPWDSRGQEATDTE
 YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTTGA
TGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC
CAACTACAACACC'TTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCTTAT
GGCCAGCGGGAATTCAGTGCAGGATTTGTTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT
GGAACATCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVIPVVYDFGDAQKTASYYSYPGQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:**Signal peptide:**

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGCAGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC
CGCCGCCCTTGTCCCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTAGCTGGTGGCCGCGCTCT
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGAGTCTCCATGTTCAAC
AGCAGCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCTTCTTCAT
ATTCGAGCGTTGGGAGTGCACCTACGTATTGGTACATTTTGTCTTCTGCAGTGCCTTCCAGCTG
TCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAAAAGAAACCTTCTGATTACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCTGGAAGAAGGAAG
GCATAGGCTTCGGTTTTCCCTCGGAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACA
TTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMENSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
 TCGCTGCTGCTTCGTGTTCTTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
 GTCCTCCCGGCTCAGAGGACCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT
 CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT
 GGCCCCGCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCGAACCACAGCCCCC
 CACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG
 GCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
 CCAACACAATGCCACAGGCCAGGGAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG
 AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG
 GAGTGGGAGAAGGTAGAACGGGGCCGGCGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG
 CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTTCAAAGTCGTCTGTG
 TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACCTAC
 CATAGTGATACCCCTACTACCCATCTGGGTGACCCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
 TGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGCCTC
 AGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
 AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGAAGTAGATCACAGGAGCACTGG
 AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG
 TCCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC
 CCTTGGTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACA
 GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
 GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCTTGTCGTGTGCTGAGCATGG
 CATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
 CAGGCCACCCCTTTCCAAAATTCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
 GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCACACTAAGGCCACAGCCCATC
 CGCGTGCTGTGTGTCCTCTTCCACCCCAACCCCTGCTGGCTCTCTGGGAGCATCCATGTCCCG
 GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTGACACGGGGTCTCCCGGATCTGGATGGCGC
 CGCCCTCTCAGCAGCGGGCAAGGGTGGGGCGGGGCCGGCCGAGAGCATGTGCTGGATCTGTTC
 TGTGTGTCTGTGTGGGTGGGGGAGGGGAGGGAAGTCTTGTGAAACCGCTGATTGCTGACTTT
 TGTGTGAAGAATCGTGTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVCPCDYNHSDTPYYPSG

Important features of the protein:**Signal peptide:**

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCCTGCGCTCGCC
CTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGCTTTCTT
CTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTCTATATCCAAGAA
ATGTTCCGATTTCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC
AGGTGAGACAGCACCTCTATGCGACTGCTGGCCTATGTTCTGGCTTGGGCTTTGGAATCATGA
GTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTTCAT
GGAGATTCTCCTCAATTCTTCCTTTATTTCAGCTTTCATGACGCTGGTCATTATCTTGCTGCATGT
ATTCTGGGGCATTGTATTTTTTGATGGCTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTTATAATCCTGGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG
AAGCCTGAACTCTGCCTGCTCTGCCAAGACAAGAACTTTCTTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCT
TTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLPGPTVGIIHGDSPOFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKWGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSCLKCLLCQDKNFLLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTTCACCAGAGTAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG
 GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC
 TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAAGTGGCTCCGGATCAGG
 GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTGATACCATTAACA
 CAGATGCTCACACTGGGGCCAGATCTGCATCTGTAAATCCTGCTGCAGGAATGACACCTGGTAC
 CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
 CAATTTTGTGCACACAACCTTGAGGCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
 TTCACGAGCCTCATCATCCATTCTTGTTCCTGGGAGGCATCCTGCCCACCAGTCAGGCAGGGGC
 TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
 GAACCCAGCAGGCCGCCTCCCAACTCCCAGTGGCACAGATGACGACTTTGCAGTGACCAACCCCT
 GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTGTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTTATC
 ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATCTTAATT
 TACCTGAAAATATTCTTGAAATTCAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAA
 CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
 TAAACATATTTGGAAAAGTGGAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTTLSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
 CTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCTGACC**ATG**
 GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCAGGCTCTCCCCAAGGCCAGCCTGC
 AGAGCTGTCTGTGGAAGTTCCAGAAAACATATGGTGGAAATTTCCCTTTATACCTGACCAAGTTGC
 CGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTACAGGGGACTCAGGCAAGGCAACTGAG
 GGCCCATTTGCTATGGATCCAGATTCTGGCTTCCTGCTGGTGACCAGGGCCCTGGACCAGAGGA
 GCAGGCAGAGTACCAGCTACAGGTCACCCCTGGAGATGCAGGATGGACATGTCTTGTGGGGTCCAC
 AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTTCTCTCAAGCCATCTAC
 AGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCCCCTTCCTCTTCCCTTGAGGCTTCAGACCG
 GGATGAGCCAGGCACAGCCAACCTCGGATCTTCGATTCCACATCCTGAGCCAGGCTCCAGCCCAGC
 CT'TCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
 ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA
 CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAAGTCTCCATCATAGAGAGCACCTGGGTGT
 CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCCATACCCGACCCACATGGCCAGGTA
 CACTGGAGTGGGGGTGATGTGCACTATCACCTGGAGAGCCATCCCCCGGGACCCCTTTGAAGTGAA
 TGCAGAGGGAAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCAGGCTGAGTACCTGCTCC
 AGGTGCGGGCTCAGAATTTCCATGGCGAGGACTATGCGGCCCTCTGGAGCTGCACGTGCTGGTG
 ATGGATGAGAATGACAACGTGCCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT
 CAGTCCACCAGGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT
 ATTCCACGTTGTGATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTC
 CAGGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGACGAGGCCAGAACAT
 CCTGCTTCTGGTGTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGTGAAG
 TCGAAGTCGCAGTACAGATATCAATGATCACGCCCTGAGTTCATCAC'TTCCAGATTGGGCCCT
 ATAAGCCTCCCTGAGGATGTGGAGCCCGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA
 CCTCGAGCCCGCCTTCCGCCTCATGGATTTTGGCATTGAGAGGGGAGACACAGAAGGGACTTTTG
 GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
 GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
 AGGCCCTGGAGCCACCGCCACGGTGAAGTGTGCTAGTGGAGAGAGTGAAGTGGTGGTGGTGGTGGT
 ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGTCCCCAGCCGGCTCTTTCCTGTGACCATC
 CAGCCCTCCGACCCCATCAGCCGAACCCCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT
 CTGCATTGAGAAATTTCTCCGGGGAGGTGCACACCGCCAGTCCCTGCAGGGCGCCAGCCTGGGG
 ACACCTACACGGTGTCTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCCTGTGCCCTCCCAA
 TACCTCTGCACACCCCGCCAAGACCATGGCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCT
 GGCCAGTGGGCACGGTCCCTACAGCTTCACCCTTGGTCCCAACCCACGGTGAACGGGATTGGC
 GCCTCCAGACTCTCAATGGTTCCCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCACGT
 GAACACATAATCCCCGTGGTGGTCAAGCACAATGCCAGATGTGGCAGCTCCTGGTTCGAGTGAT
 CGTGTGTGCTGCAACGTGGAGGGGCAAGTGCATGCCAAGGTGGGCCGATGAAGGGCATGCCCA
 CGAAGCTGTGGCAGTGGGCATCCTTGTAGGCACCCTGGTAGCAATAGGAATCTTCTCATCCTC
 ATTTTCAACCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCT
 GAAGGCGACTGTCT**TGA**ATGGCCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAG
 TCCCTTGGGAGAGAGCCAGCACCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCTCCA
 TCTGCCCTGGGGTGGAGGCACCATCACCATCACAGGCATGTCTGCAGAGCCTGGACACCACTT
 TATGGACTGCCCATGGGAGTGCTCCAAATGTAGGGTGTTTGGCCAAATAATAAAGCCCCAGAGAA
 CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSPQALPKAQAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKMDGDAQSGHQATATVEVSI IESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVLPLRAGQNILLVLAMDLAGAEGGFSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTADADLEPAFRIMDFAIERGDTEGTFGLDWEFDSGHVRLRLCKNLSY
EAAPSHEVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALT LAPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTALHWVEP
REHIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTTLVAIGIFLI
LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCACGCCTG
 AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA
 TCTCTTCACTTCAAGTCCCCTTTCTCAAGAATCCTCTGTCTTTGCCCCTCTAAAGTCTTGGTAC
 ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
 TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATCCAATGAGACTAGCACC
 TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
 TGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA
 GCATAGTCACCAACTCTGAGTTCATACAACTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
 TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
 GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGGGGCCAGCACAGTCCACCACT
 CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
 AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC
 CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
 CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
 GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
 AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCA
 GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG
 TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
 GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACT
 CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCTCC
 AGTGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCAC
 CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
 CCTCCAGTGGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA
 GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG
 CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA
 ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
 ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGTGAGGCAAAGCCTGGTGGGTCCCTGGT
 GCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTGTGGCGGCCGTGGGGCTCTTTGCTGGGC
 TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT
 CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCACAGGCCAG
 GTGGAGTCCTAACCTGTTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA
 ACAGCGGGCCCTGAGCAGCCCCGGAAGCAAGTGCCGCAATTCCTCAGGAAGGAAGAGACCTGGGCA
 CCCAAGACCTGGTTTCCTTTTATTTCATCCCAGGAGACCCCTCCCAGCTTTGTTTGAGATCCTGAA
 AATCTTGAAGAAGGTATTCTTACCTTTCTTGCCCTTACCAGACACTGGAAAGAGAATACTATAT
 TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTG
 CCCCAGGGTGGGTATCTAGCTCTGAGATGAAGTATAGGAGAAAACCTCATGCTGGACTC
 CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVSSGASTATNSGSSVTSSGVSTATISGS
 SVTSNGVSIIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
 GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
 TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES
 STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS
 GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
 TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES
 STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA
 AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFWRRPVSSI
 AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTACAGGTG
GGAGGCTGGAACATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT
GGTGTGATGGAAGGGATTTACTTTTATACTGACTCTGTTTTGGGGAAGCTTTTTTGAAGCATTT
TCATGCTGAGTCCCTTTTTACCTTTGATGTTTTGTAAACCCATCTTGGTATCGCTGGATCAACAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAGT
GATTATAACTGGGGATGCATTTGTTCTGGAGAAAGAAGTGTCATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGGTTGGGCCATGCAGGCTGCTGCCTATAT
CTTCATTGATAGGAAATGGAAGGATGACAAGAGCCATTTGGAAGACATGATTGATTACTTTTGTG
ATATTCACGAACCACTTCAACTCCTCATATTCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAACATTCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT
CCCAGGGAAATCCACTTTCACGTCCACCGGTATCCAATAGACACCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAG
GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTCAATCCACCTTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTCTATACTGTATTGGACCTGTTGAGCCCTGCAATGTGCCTACTCAT
ATATTTGTACAGTCTTGTTAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
GAATATTTGGTGGACTGGAGATCATAGAATTGCGATGTTACCGACTTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTG
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTGCATGACTATGTCGAATATTTCTTACT
GCCATCATTATTTGTTAAAGATATTTTGCACTTAATTTTGTGGGAAAAATATTGCTACAATTTT
TTTAATCTCTGAATGTAATTTGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT
GGGCCAGAATATTATTAAACAATCATCAGGCTTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIYFILTLFWGSFFGSIFMLS
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCIMRYSYLRLKICLKASLKGVPFGFWAMQAAAYIFIHRKWDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSKSRNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLLOGDFPREIHFHVHRYPIDTLPTSKEDQLWCHKRWEEKEERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVLVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIEIACRYRLHKKPHLNSKKNE

Important features of the protein:**Signal peptide:**

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
 TCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTTCATAGTGTGAGATCAACCCACAGGAATA
 TCCATGGCTTTTGTGCTCATTGTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT
 CACTGGACCGGCAAGTTTGTCCAGGCCCTTGGTGGGGGAGGACGCCGTGTTCTCTGCTCCCTCT
 TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCGGTTCTTCAGGAATCAGTTCATGCTGTGGTCT
 CACCTCTACAGAGATGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAAGTGA
 GTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTCTAAGGCTAAAAACATCACTCCCTCGG
 ACATCGGCCTGTATGGGTGCTGGTTTCACTTCCAGATTTACGATGAGGAGGCCACCTGGGAGCTG
 CGGGTGGCAGCACTGGGCTCACTTCTCTCATTTCATCGTGGGATATGTTGACGGAGGTATCCA
 GTTACTCTGCCTGCCTCAGGCTGGTTCCCCCAGCCCCACGCCAAGTGGAAAGGTCCACAAGGAC
 AGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
 ATTTAGTCCAGGAAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCAATGA
 GGTGGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTCTTA
 TTTTACTCGGGTTACTCTGTGTGCCCTGTGTGGTGTGTGTCATGGGGATGATAATTGTTTTCTTC
 AAATCCAAAGGGAAAAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACGGACAGGCAGAAATTGAG
 AGACGCCCCGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCG
 TTTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA
 TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
 CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTGCGGATGACGTAGACAGGGGGAAGAACA
 ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCCTCAGACTGACAACAGAACATTTGTATTTT
 ACATTCATCCCCATTTTATCAGCCTCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCCT
 GGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCTGC
 TGACATGTGAGTTTGAAGGCTTGTGAGACCTTATATCCAGCATGCGATGTATGACGAGGAAAAG
 GGGACTCCCATATTCATATGTCCAGTGTCTGGGGATGAGACAGAGAAGACCCTGCTTAAAGGGC
 CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGTCTCCGACAGGTGGCCCCAGCTTCTCT
 CCGGAGCCTGCGCACAGAGAGTACGCCCCCCTCTCTCTTATAGGGAGCTGAGGTTCTTCTGCCC
 TGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAG
 TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAA
 CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTGCTAGTCACGG
 ACAGTGATTCTGCTCAGAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT
 TGAGGGCACAGTGTGCTAATGATGTGTTTTTATATTATACATTTCCACCATAAACTCTGTT
 TGCTTATTCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACACC
 TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTTCACTGATTCTATAAGCCCAGCAT
 TACCTGATACCAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACACAGGTCCATATCC
 CTCATTAACACAGACACAAAAATTTAAATAAAATTTTAAACAAATTAACTAAACAATATATTTA
 AAGATGATATATACTACTCAGTGTGGTTTGTCCACAAATGCAGAGTTGGTTAATATTTAAAT
 ATCAACCAGTGTAATTCAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
 LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
 VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGDLSSDSRANADGYSLYDVEISI
 IVQENAGSILCSIHAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVMGMIIVFFK
 SKGKIQAELDWRRKHGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
 TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT
 FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLTLCQFEGLLRPYIQHAMYDEEKG
 TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
 TTTGTTGGGAACCTGGGTATCGGCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA
 CTCACTGTTCATTATGTGAGATATAATCAAAGAAGACCTACAATTACTATAGCACATGTTCATTACAAAC
 TGACAAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAATTTTACAGAAATGAGCCAGAGACTTGAAT
 CAATGGTGAAAAATGCATTTTATAAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC
 AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTTCACTCTACTGAGGATCCTGA
 AACTGTAGATAAAATGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCCATAAGTAG
 ATCCTCACTCAGTTAAAATTAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA
 ACACGAAGAAGTAAACTCTAGGTGAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATG
 GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAATGCCACATGGC
 TTGTGAGTGTCTCACTGTTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACA
 ATAAACCTTCGAAATGAAACGGGCTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACA
 TGACTATGATATTTCTCTGCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCTC
 TCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTGTGACAGGATTTGGAGCACTGAAAAAT
 GATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC
 TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT
 GCCAGGGTGACTCTGGAGGACCCTGGTGTAGTTCAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG
 AGCTGGGGAGATGAATGTGCGAAACCAACAAGCCTGGTGTTTATACTAGAGTTACGGCCTTGCGGGACTG
 GATTACTTCAAAACTGGTATCTTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTGTTTTTGT
 GGTGTGGAGGCCATTTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTGTGACTGATCTCA
 ATAAACTGTTTGCTTGATGCATGTATTTCTTCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCA
 GATCAACTCTGTCTGTGAGCAATAGTTGAAACTTTATGTACATAGAGAAATAGATAATACATATTAC
 ATTACAGCCTGTATTCATTTGTTCTCTAGAAGTTTGTGCAATTTTGACTTGTGACATAAATTTGTAAT
 GCATATATACAATTTGAAGCACTCCTTTTCTCAGTTCCTCAGCTCCTCTCATTTACAGCAAATATCCATTT
 TCAAGGTGCAGAACAAGGAGTGAAAGAAAATATAAGAAGAAAAAATCCCCTACATTTTATTGGCACAGAA
 AAGTATTAGGTGTTTTCTTAGTGGAATATTAGAAATGATCATATTTCATTATGAAAGGTCAAGCAAAGACA
 GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
 ATATATCCTTATTTTCATTTCCAAACAACACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTGTGACCT
 ATAATAATTATACAACTTCATGCAATGTACTTGTTCTAAGCAAATTAAAGCAAATATTTATTTAACATTG
 TTACTGAGGATGTCAACATATAACAATAAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTNYYSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHPSHDYDISLAELSSPVPYTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSGGPLVSSDARDIWYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCCTG
 CCCCAGATGAGCCCCGCGTCCCGACTATCCCCAGGCGGGCGTGGGGCACCGGGCCAGC
 GCCGACGATCGCTGCCGTTTTGCCCCTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT
 ACGGGGCTCACAAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCGTCTACGCCCTCAA
 TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTGGCAGTTTCTGCTTGGATGAGGGACTACCTAA
 ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTTGACTTACTTTCCCT
 GTGGTTCATCCGGTCATGATTGCTGTTTGTCTGTTTCCCTTATCATTGTGGGGATGTTAGGATATTG
 TGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCAATGGTACTTTGGAAGTTTGTCTGTCAATTTCT
 GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTATGGTTCAGTACAATGGTCA
 GATATGGTCACCTTGAAAGCCAGGATGACAAATATGGATTACCTAGATATCGGTGGCTTACTCA
 TGCTTGGAAATTTTTTTCAGAGAGAGTTTAAAGTGTGTGGAGTAGTATATTTCACTGACTGGTTGG
 AAATGACAGAGATGGACTGGCCCCCAGATTCCCTGCTGTGTGTAGAGAATTCCAGGATGTTCCAAA
 CAGGGCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTTGTGGGAAGAAAATGTATTCCTT
 TTTGAGAGGAACCAACAACCTGCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA
 TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG
 ACAGACCAAAATGATGTCCTTGAAGAATGACAACCTCTCAGCACCTGTCATGTCCTCAGTAGAACT
 GTTGAAACCAAGCCTGTCAAGAATCTTGAACACACATCCATGGCAAACAGCTTTAATACACACT
 TTGAGATGGAGGAGTTAATAAAAGAAATGTCACAGAAGAAAACCAAACTTGTATTATTGGACT
 TGTGAATTTTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAAATAAAATGTTGCCATAAAA
 TAACACCTAAGCATATACTATTCTATGCTTTAAATGAGGATGGAAAAGTTTCATGTCATAAGTC
 ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC
 TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCCGCA
 TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA
 CTAGTATATAAAGTACTAATTAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA
 CTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACCTTCAATATTGGTGACT
 ACCTAAATGTGATTTTTGCTGGTTACTAAAATATTCTTACCCTTAAAAGAGCAAGCTAACACAT
 TGTCTTAAGCTGATCAGGGATTTTTTGTATATAAGCTGTGTAAATCTGTATAATTCAGTCGAT
 TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAAATTTGTCCTGTATAGCATCATT
 ATTTTTAGCCTTTCCCTGTTAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAAC
 TGTTATTTAAATACCTAACCACTAATTTTGAAAATTACCAGTGTGATACATAGGAATCATTATTC
 AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTGAGA
 AAGGACTTGTATGCTGTTTTCTCCCAAATGAAGACTCTTTTTGACACTAAACACTTTTTAAAAA
 GCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATCTACAGAAAA
 TAGTGTCTTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAAAACATGTGACAAATTTAGAGATT
 CTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAATTTTTTTACAA
 GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCAATTTACTGTATTTTGTGTATTTTGTATTAT
 TTCTCAGAATATGGAAGAAAAATTAATGTGTCAATAAATATTTTCTAGAGAGTAA

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCTTCTCTGATAAAGCCCTACCAGTGC
 GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCCTTAAAAAAGTCTTGAAAGAGAAGGGACAAAGGAACA
 CCAGTATTAAGAGGATTTTCCAGTGTCTTGGCAGTTGGTCCAGAAGGATGCCCTCCATTCTGCTTCTCACCTG
 CCTCTTCATCACAGGCACCTCCGTGTCAACCCGTGGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGC
 CCTGGAGGAACACTGACCACCAGTTGGATGAGTCTCAAGGTCTCCTCTATGTGACAACCATGTGAATGGGGAG
 TGGTACCCTTACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
 CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAAGGCGACGGCATTGTGCAACGCCAGGCTTGTGCCAGCT
 TCAATGGGAAGTGTGTCTCTGGAACACCACGGTGGAAGTCAAGGCTTGCCCTGGAGGCTACTATGTGTATCGT
 CTGACCAAGCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTATGACATCTGCGACGAGGACTGCCATGG
 CAGCTGCTCAGATACCAGCGAGTGCACATGCGCTCCAGGAAGTGTGCTAGGCCCTGACAGGCAGACATGCTTTG
 ATGAAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTCTACCCTGT
 GAGTGTGGGGTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
 TGGTGGCTGCAGCCACTCTTGCCTTGGATCTGAGAAAGGCTACCAGTGTGAATGTCCCCGGGGCCTGGTGCTGT
 CTGAGGATAACCACACTTGCCAAGTCCCTGTGTGTGCAAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
 CTGGTTGGTGGCCTGGAGCTCTTCTGACCAACACCTCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACAT
 CCTCTTCTCTCAAGACATGTGTACAGTGGTGCATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA
 CAGGTCTACCCAAGCAGACCCCGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTG
 ACCTGCGAGTTTCCACGCCGTGTACACCATTTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAT
 CATGAGCCGAAATCATGGGATCTTCCCATTCACCTCTGGAGATCTTCAAGGACAATGACTTTGAAGAGCCTTACC
 GGGAAAGCTCTGCCCACCCTCAAGCTTCGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC
 TTGGAAAGCTTGGTGGAGAGCTGCTTTGCCACCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCAT
 CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGG
 TCCCTGTCTTCAAGTTTGTGGGCAAGACCACAAGGAAGTGTCTGCACTGCCGGGTCTTGTCTGTGGAGTG
 TTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGGCGAATGCGTCGTGGGGCAGGAGGAGAGCACTCAGC
 CGGTCTACAGGGCCAGACGCTAACAGGCGGCCGATCCGCATCGACTGGGAGGACTAGTTCGTAGCCATACCTC
 GAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACC GCCCTCTAAGAACATCTGCCAACAGC
 TGGGTTCAAGACTTCACACTGTGAGTTCAGACTCCAGCACCAACTCACTCTGATTCTGGTCCATTAGTGGGCA
 CAGGTACAGCACTGCTGAACAATGTGGCCTGGGTGGGGTTTCATCTTTCTAGGGTTGAAAATAAAGTGTCCA
 CCCAGAAAAGACACTCACCCATTTCCCTCATTTCTTCTACACTTAATACCTCGTGTATGGTGAATCAGAC
 CACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCCTAGAAAAATTAAACAGTTACTGAAATTATGA
 CTTAAATACCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAAATTTCAATTCAAATGCAGACTAA
 TTATAGGGAATTTGGAAGTGTATCAATAAACAGTATATAATTTT

FIGURE 110

MPFFLLLTCLFITGTSVSEFVALDPCSAYISLNEPWRNTDHLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
 TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCWNTTVEVKACPGGYVYRLTKPSVCFHV
 YCGHFYDIEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
 LRSDGKTCEDVEGCHNNNGGCSHSLGSEKGYQCECPRGLVLEDNHTCQVPVLCKSNAIEVNIIPRELVG
 LELFLTNTSCRGVSNHTVNILFSLKTCGTVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
 CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEFPYREALPTLKLRLDSLYFGIEPVVHV
 SGLESLVESCFATPTS KIDEVLKYLLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRV
 LVCVGLDERSRCAQGGCHRRMRGAGGEDSAGLQGQTLTGGPIRIDWED

Important features of the protein:**Signal peptide:**

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

GAGAGAGGCAGCAGCTTGTCTCAGCGGCAAGGATGCTGGGCGCTGAGGGACCAAGGCCCTGCCCTGCACCTCGG
GCCTCCTCCAGCCAGTGTGTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT
CCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
GGTTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCCCTCTGCGGGCAGCCTCTCCACTTCATCCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
CCCGAAGGGCCTGCAGTGGCAGTCCGCCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGGAACCTGGTTCTCTGCCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGCATGCGGAACCTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCCTGTCTTGCCTG
TGGGAAGAGCCTGAAGACCCCCGTGTGGTGGGTGGGGAGGAGGCCTCTGTGGATTCTTGGCCTTGGCAGG
TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCTGGACCCCCACTGGGTCTCTACGGCA
GCCCCACTGCTTCAGGAAACATACCAGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAG
CTTCCCATCCCTGGCTGTGGCCAAGATCATCATATTGAATTCAACCCCATGTACCCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTT
GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGAATGGAGGGAA
GATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCATTGACAGCACACGGTGCAATGCAGACGATGCGT
ACCAGGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGAAGGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCTGTATGTACCAATCTGACCACTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG
CTGCGGGGGGCCGAGCACCCCAAGGATATACACCAAGGTCTCAGCCATATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGCTGTCTGCCCCCTTTGCACTGCTGGGAGCCGCTTCCTTCTGCCCCGCCCCACCT
GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCCTTGGGTACACCCCTCTGCCACAGCCTCAGCAT
TTCTTGGAGCAGCAAAGGGCCTCAATTCCCTGTAAGAGACCCCTCGCAGCCCAGAGGCGCCCAGAGGAAAGTCA
GCAGCCCTAGCTCGGCCACACTTGGTGCTCCCAAGCATCCAGGGAGAGACACAGCCCACTGAACAAGGTCT
CAGGGGTATTGCTAAGCCAAGAAGGAACTTTCCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAAAGCC
CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAA
GCCTACTAGAGCAAGAAACCAGTTGTGAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
ACTGTTGTCTATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA
AAAA

FIGURE 112

MLQDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALSLASIIIVVVLIKVILDKYYFLCG
 QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
 FTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
 KTPRVVGEEASVDSWPWQVSIQYDKQHVC GGSILDPHWLTAAHCFRKHTDVFNWKVRAGSDKL
 GSFP SLAVAKIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIIGWG
 FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCCQDSSGGPLMYQS
 DQWHVVGIVSWG YCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRD

Signal peptide:
amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCCT
GTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG
CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCTGAATTTT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACTTTCTTTGAACAGTCAGTG
ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTCTTAAAAATTCCAAAATCTGGA
GATTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG
ACTTTGAGGAGGAGGAGAAGATCTTCACTTTCCTGCCAACGAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA
AGAACTTCCAATAAATGACTATACTGAAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAG
GTTATTGTTGTATTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCTCTGTGAACCTTTACTA
GGCTACTACCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTG
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCCT
GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPPKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWPVPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIIYCRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCGGCA
GGGCGGGCGGCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGTTCTCTCTGCCATCCTGGGGCT
GGCCGGCTGCATCGCGGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCC
TATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCCT
GGGTGCCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTG
TCTGTGTTTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG
GATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGCAGGCTCTGTTTCGTGGGCTGGGTGCTGGAGGCC
TCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCTGGCACCAGAAGAAACCACTACAAA
GCCGTTTCTTATCATGCCCTCAGGCCACAGTGTTCCTTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTT
TGGGTCCAAACACCAAAAACAAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCACGACTATGTTAATGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA
CCCAAAAACAAGGAGATCCCATCTAGATTCTTCTTCTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCT
CGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCACCATAAAACA
GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTC
TTCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAAACTTTGAAAGGAAA
GAGTAGACCCAAAGATGTTATTTTCTGCTGTTGAATTTTGTCTCCCCACCCCAACTGGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTT
ACACTGTGATCTTAAAGTTACCAAACCAAGTCATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTG
TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTTTCTGT
CGCGGGTCAGAAATGTCCCTAGATGAATGAGAAAATTATTTTAAAAATTTAAGTCCTAAATATAGTTAA
AATAAATAATGTTTGTAGTAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTC
ATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCAGAGT
TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGCTCTACAAAATACAGAGAGAAAAATCAGCCAGTCA
TGGTGGCATAACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCAGGGAGGT
TGGGGCTGCAGTGACCATGATCACACCCTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAAATAAATAATGGAACACAGCAAGTCTAGGAAGTAGGTTAAACTAATCTTTAA

FIGURE 119

GGAAAACTGTTCTCTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCGGAGTCC
AGCTGGCTAAAACTCATCCCAGAGGATAATGGCAACCCATGCC'TAGAAATCGCTGGGCTGTTTCTTG
GTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTCAATGCCCTAGTGGAGAGTGTCCGCCCTTCATT
GAAAAACAACATCGTGGTTTTTGGAAACTTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTTCTCCGGACCTACAGGCAGCCAGAG
GACTGATGTGTGCTGCTTCCGTGATGTCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
CATCACGGGCATGGTGGTGCTCATCCCTGTGAGCTGGGTGCCAATGCCATCATCAGAGATTTCTATA
ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCA
CTGGTGCTGATGTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTGTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACCTTCCCATCGACAACCCAAAAAGTTATCACACCGGAAAGAAGTACCGAGCG
TCTACTCCAGAAGTCAGTATGTGTAGTTGTGTATGTTTTTAACTTTACTATAAAGCCATGCAATG
ACAAAAATCTATATTACTTTCTCAAAATGGACCCCAAAGAACTTTGATTTACTGTTCTTAACTGCCT
AATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAAGCTATTTAGCAGAATGAGATA
TTAAACCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
CTCTTTTATCATTTACTTCAAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAACTTTATATCTCACATAGAGACATGCTTATATGGT
TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAA
ATCATGGATAGGGTTGAAGAAGGTTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCCTCCA
TTTATAATGAAGATTAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATAT
GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCAGAGGCTTTTTTTTT
CTTGTGTATTAAATTAACATTTTTTAAACGCAGATATTTGTCAAGGGGCTTTGCATTCAAAGTGCCT
TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGTGGTTTTAGGAAAGTG
AAAATATTTTTGTTTTGTATTGTTGAAGAAGATGATGCATTTTGACAAGAAATCATATATGTATGGAT
ATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAGAGCAGAAAAATA
TGTCTTGGTTTTTCATTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTTGAGAACTTCACCT
GCTCCTATGTGGGTACCTGAGTCAAAATGTCAATTTTGTCTGTGAAAAATAAATTTCCCTTCTTGTA
CCATTTCTGTTTAGTTTTTACTAAAATCTGTAAATCTGTATTTTCTGTTTATCCAAATTTGATGAA
ACTGACAATCCAATTTGAAAGTTTGTGTGACGCTGTCTAGCTTAAATGAATGTGTTCTATTGCTT
TATACATTTATATTAATAAATTGTACATTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQCK
IYDSLLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLI VGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGCGGAG
CCAGACGCTGACCACGTTCCCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCCGCGAGCC
GGGAGCCATGCGACCCAGGGCCCCGCGCCTCCCCGAGCGGCTCCGCGGCCTCCTGCTGCTCC
TGCTGCTGCAGCTGCCCCGCGCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCGGGTACACCTGGGATCCCAGGTCGGGATG
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCTGGACACCCAACTAC
AAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGATACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTCAAGTGGCTCACTTCGGCTAAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTCAAGACCTCTTCCC
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATCAACAATTAATATTCATCG
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGG
TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC
ATTATGAAGAACTACCAAAATAATGCTTTAATTTTCATTTGCTACCTCTTTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTTACAGACCAAAGTGTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTG
CTTCAATCAAAAGTGGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT
CTCTCAACCTATAATTTGGAATATGTTGTGGTCTTTTGTTTTTCTCTTAGTATAGCATTTTAA
AAAAATATAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGT
TAAATAAAAATTATTTCCAACA

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
 CTGGGTGCTCATCACGGGAACCTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
 ATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATGTAGGATCCAGTTTTTTTTTTTA
 ACCGCCCCCTCCCCACCCCCCAAAAACTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC
 TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGGAG
 TGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAGGGGTCCAATTTT
 TCTTCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTCATGCAACTG
 GCCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAAAGGATGGGTTTCAATG
 TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCCACTGTCTTACTGACAATG
 CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA
 ATCTCAGAAATTACAGGAGATACCCTCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT
 ATAACAGCCTTCAAAAACCTTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACCTGGCTATAC
 CTTGACCATAACCATATCAGCAATATTGACGAAAAATGCTTTTAAATGGAATACGCAGACTCAAAGA
 GCTGATTCTTAGTTCCAATAGAAATCTCCTATTTCTTAACAATACCTTCAGACCTGTGACAAATT
 TACGGAACTTGGATCTGTCCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTG
 CGGAAGCTGCTGAGTTTACATTTACGGTCTAATCCCTGAGAACCATCCCTGTGCGAATATTCCA
 AGACTGCCGCAACCTGGAACTTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG
 TCTTTGCTGGCATGATCAGACTCAAAGAACTTCACCTGGAGCACAATCAATTTTCCAAGCTCAAC
 CTGGCCCTTTTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGTCAGTGGGAATAAATCAGTGT
 CATAGGACAGACCATGTCCTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGA
 TCGAAGCTTTTCACTGGACCCAGTGTTTTCCAGTGTGTCCCGAATCTGCAGCGCTCAACCTGGAT
 TCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAG
 TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTTGTAACCTGGCTGAAAAGTT
 TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
 ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGATCTGGCCAG
 GGCTCTCCCAAAGCCGACGTTTAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTT
 TGCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT
 TTCCATAAAATCATCGCGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCCTGCTGGTTAT
 CTACGTGTCATGGAAGCGGTACCTGCGAGCATGAAGCAGCTGCAGCAGCGTCCCTCATGCGAA
 GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCAGCACCCAGGAATTTTATGTA
 GATTATAAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCTGCACCTA
 TAACAAATCGGGCTCCAGGGAGTGTGAGGTATCAACCATTGTGATAAAAAGAGCTCTTAAAGCT
 GGGAAATAAGTGGTGTCTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCTCCCC
 TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCTTGTCCGTTTTAGTGCAATCATAATACT
 GGTCATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT
 GAACTCCGGTTTAAATATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTGCGATTT
 GTTTTAAGATAAAACTTCTTTTATAGGTAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTPLLTLSSAERGCPKGCRCGKMVYCESQKLQEIPSSISAGCLG
LSLRYSNLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDSLQNLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRI FQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLEHNQFSKLNALFPRLVSLQNLQWNKISVIGQTMSTWSSLQRLDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNKLTFIQEILDSWISLNDISLAGNIWECNRNCSLVN
WLKSFKGLRENTIICASPKEQGVNVIDAVKNYSICGKSTTERFDLALALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMQLQQR
SLMRHRHKKRQSLKQMTPTSTQEFYVDYKPTNTETSEMLLNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCGCTACTGCTGAATGTCCGTCCCGGAGGAGGAGGAGGCTTTTGCCGCTG
 ACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGCCGAGCT
 AGCAACCTTTCCCTGGATCTCACAAAACCTCGACTCCAAATGCAAGGAGAAGCAGCTCTTGCTC
 GGTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACAGCCCTAGGGATC
 ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACACGTAGT
 GTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTGGCAAAAGTGAAG
 ATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTTATTGGCCAGTTT
 TTAGCCAATCCAACCTGACCTAGTGAAGGTTTCAGATGCAAATGGAAGGAAAAAGGAAACTGGAAGG
 AAAACCATTGCGATTTTCGTGGTGACATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC
 GAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA
 ACCACTTATGATACAGTGAAACACTACTTGGTATTGAATACACCCTTGAGGACAATATCATGAC
 TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA
 TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCG
 ACTGACTGCTTGATTGAGGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTACC
 ATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGTTCTGGCTTACTTATGAAAAAATCAGAGAGA
 TGAGTGGAGTCAGTCCATTTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGIIIEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMOMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCGLVASILGTPADVIKSRIMNP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTCGGCGGGCGGCGGAGAGCGGCCGGGCGTCAGCTCCTCGAC
 CCCCCTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGGCCAGGCCCGGCATGG
 AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCTCGGGGGGCATCGGCGCGGCGTGGCC
 CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA
 GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCCACAGATGTGACCTAT
 CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
 TGCATCAACAATGCTGGCTTGGCCCGGCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAA
 GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
 AGGAGCGGAATGTGGACGATGGGCACATCATTAAACATCAATAGCATGTCTGGCCACCGAGTGTTA
 CCCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCAGTGCCTGACAGAGGGACT
 GAGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG
 AGACACAATTGCGCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA
 ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCCTCAGCACCCCGCACA
 CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC
 TCCTTCCCTCCCCACCCTTCATGGCTTGCTCCTGCTCTGGATTTTAGGTGTTGATTTCTGGAT
 CACGGGATACCACTTCTGTCCACACCCGACCAGGGGCTAGAAAATTTGTTTGAATTTTATA
 TCATCTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC
 TAATTGTTTTACTTGTAACTTGTCTTGTGCCCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTG
 TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTTGACCTCAACGTCTG
 TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTACCTTATATCTGTGTTGTATCCAGGGCTCC
 AGACTTCCTCCTGCTGCCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC
 AGCCCAGTCTTGGCTTCTTGTCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG
 CAGAACACCAGGGCCTGGCCCAGTGGATTTTCATGGTGATCATTAAAAAAGAAAAATCGCAACCAA
 AAAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGA A VARALVQQGLKVVG CARTVGNIEELAAECKSAGYPGTLI
 PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPD TLLSGSTSGWKDMFNVNVLALSICTR
 EAYQSMKERNVDDGHIININMSGHRVLP LSVTHFY SATKYAVTALTEGLRQELREAQTHIRATC
 ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
 199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

TTGGGCGCTGCTGTTGCTGAC

FIGURE 130

MGLLLLVLFLLPVAITIMSLPPSFDCGPFCRVSVAREHLPSRGSLLRGPRPRIPLVSCQPV
KGGHTLGESPMFPKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
 ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
 TCCATCTGTGTGTCGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTCTGACATCCATTCCAACAG
 GAATACCAGAGGATGCTACAACCTCTACCTTCAGAACAACCAAATAAATAATGCTGGGATTCTTCAGAT
 TTGAAAACTTGCTGAAAGTAGAAGAATATACCTATACCACAACAGTTTAGATGAATTTCTTACCAACCT
 CCCAAAGTATGTAAAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAA
 AAATTCCTATCTGGAAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
 TTCCGAGACAGCAACTATCTCCGACTGCTTTTCTGTCCCGTAATCACCTTAGCACAAATCCCTGGGGTTT
 GCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCATCTCTTCAAG
 GTCTCACTAGTCTAAAACGCTGGTTCTAGATGGAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTT
 TTCTTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGCGGAATTCCTGACTGCTGCACCAGTAAACCT
 TCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCCAAATGCTTTT
 CTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTT
 GATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
 GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA
 AGGTTTCGTGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAAGC
 ACCATTTCAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCCAAGGACAGTGGCCAGCTCCAGTGAC
 CAAACAGCCAGATATTAAGAACCCCCAAGCTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAA
 CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTATATCTCTTGGAACTTGCTCTACCTATG
 ACTGCTTTGAGACTCAGCTGGCTTAACTGGGCCATAGCCCGGCATTTGGATCTATAACAGAAACAATTGT
 AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTCACCTATAAAGTATGCATGGTTC
 CCATGGAAACCAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAACTGCACCCCTT
 CGAATGTACAACCCTACAACCACCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACC
 TTTGGCTGCCATCATTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTTGGTATG
 TTCATAGGAATGGATCGCTCTTCTCAAGGAAGTGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
 GCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAATCAGGGAACTTCTTTTCAGATGTTACCAAT
 AAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAAACACACCATATTTCTCCTAATGGAATGAATCTGT
 ACAAAAACAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCAGACTCAGATCAC
 TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTTTGGGTTTTTAAACCTAAGGGAGGTGATG
 GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCND RFLTSIPTGIPEDATTLYL
 QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTTITYDSLSKIPYL
 EELHLDNDSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
 QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
 RVPPNAFSYLRQLYRLDMSNNNLSNLPQGI FDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV
 KVNVRGLMCQAPEKVRGMAIKDLNAELFDC KDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD
 IKNPKLTKDQQTGSPSRKTITITVKS VTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET
 IVTGERSEYLVTALEPDSPIKVCMPMETSNLYLFDETPVC IETETAPLRMYNPTTTLNREQEKE
 PYKNPNLPLAAIIGGAVALVTIAL LALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
 ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLYKNNHSESSSNRSYRDSGIPDSDHSHS

Important features of the protein:**Signal peptide:**

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645**Amidation site.**

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

CCGTCATCCCCCTGCAGCACCCCTTCCCAAGAGTCCTTTGGCCAGGACCCCAAGGCTTCTTGCCA
GCCCTGCCGGGCCACTTGTCTTCCATGTCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGGCTTGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
GCGCTGGAGAGAGCCACCGTCTTCTTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGGCCCAGGAGCCCCTGC
TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCAGCTGACCCCTCCAGCCCCG
GTTTTGGAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTACCCACGTTCCGGC
CCCAGGACTCATTTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCACTGCTGGGAACCGGGACG
GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAAGGAGCCTCATGACCAAGCCCGGCTGCTC
AGGCTACTGCCTGTCCCACCAACTGCTCTTCTTCTCCTGCGCCAGAATGAGGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCTACCCTACCCGGGACATCTTCATGGAAAACATCATGTTCTG
TGGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTTTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGC
AATTCTCTTGCTCATCTCCCGAGTAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAAT
TTTTATATTTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
CTCAAGAGATCCGCCCACCTCAGGCTCCCAAAGTGTGGGATTATAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC
ATGGGGGCTCTCTCCCTTAGATGGCTGCTCCTCCACAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCTTATACATCTTGCGAGAATACCCCCAGCAAACAGAGACCACCCATCCACACCG
CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCC
TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCCGACACCAATCTTGGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAA

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGCTG
CTGCTGCTATCGGGGATGTCCAGAGCTCGGAGGTGCCCCGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGCGTGCAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGGCGGCCCCGAGTGTGGTAGACGGAGAAGAGCACGTCGGTTTCCTTAAG
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGGAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT
CCACCTTCTTACTTTATTAAAAGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCAATGGT
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTGACAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGGAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTGATCTCTTACAACGTGTATGTT
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGTAT
GAGGTCAATATTGATGTCACCTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAATTAT
ATGAACTACTATACATTATGTATATTAATTAACATCTTAATCCAGAAATCAAAAAAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLLSGDVQSSEVPGAAAEGSGSGVGIGDRFKIEGRAVVPGVKPDWISAA
 RVLVDGEEHVGFLKTDGSEFVVHDIPSGSYVVEVVSAPYRFDPPVRVDITSKGKMRARYVNYIKTSE
 VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMVLPLLI FVLLPKVVNTSDPDMRREME
 QSMNMLNSNHELDPDVSEFMTRLFSSKSSGKSSSGSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGGGA
 CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAAACAA
 GTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
 CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
 CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAGGCTTGGCAGT
 TTTTCTTACTCCTGTGGTCTCCAGATTTAGGCCTAAGATGAAAGCCTCTAGTCTTGCCTTCAGC
 CTTCTCTCTGCTGCGTTTTATCTCTATGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
 AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG
 TGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA
 AAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTCTATCTGGACAGGGTATTTAA
 AAATACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCCTTTCTTA
 CCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
 AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
 TTTGGGGGAACTAGACATTCTTCTGCAATGGATGGAGGAGACAGAATTAGGAGGAAAGTGATGCTG
 CTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA
 ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
 CTTCTTGATGATTGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT
 TGTAATATCTTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTGTATTTA
 ATGTATTTATTTTTTACTTGGACATGAACTTTAAAAAAATTACAGATTATATTTATAACCTG
 ACTAGAGCAGGTGATGTATTTTTATACAGTAAAAAAAACCTTGTAATTCTAGAAGAGTGG
 CTAGGGGGGTATTTCATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
 ATTTGAAATTGAACCAATGACTACTTAGGATGGGTGTGGAATAAGTTTGTGTGGAATTGCAC
 ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAG
 CCAGGAATCCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTCTTTGCATACCAAAAAAA
 AAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPMSFSIYSLQVPAVPG LTCWALTAEPGWGQNKGATTCATNSHSDSELRPEIF
 SSREAWQFFLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
 FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
 SLANSFLT IKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
 E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCGCCATCCAGCC
 TAGCGTGTCCACGATGCGGCTGGGCTCCGGGACTTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
 TAGGGATCGCGGTCTTCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCTCTGCCAGAGCGGAA
 CACGGAGCGGAGCCCCAGCGCCCGAACCTCGGCTGGAGCCAGTTCTAACTGGACCAGCTGCC
 ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGATTTTGTGTTTG
 GGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTTACCTTGTGGAAAAAGGAGCATCTCACAGT
 TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
 CCTTCTGGCTTTGTGACGTCATCAGGAACCTCAATTCTCTGCACTGCTGGAAGACAGTGTGA
 TAAGACAAGCAAAAGCAGCTGGAAAAAGAATAGTCTTTTATGGAGATGAAACCTGGGTAAATTA
 TTCCCAAAGCATTTTGTGGAATATGATGGAACAACCTCATTTTTCGTGTCAGATTACACAGAGGT
 GGATAATAATGTCACGAGGCATTTGGATAAAGTATTAAAAAGAGGAGATTGGGACATATTAATCC
 TCCACTACCTGGGGCTGGACCACATTGGCCACATTTACAGGGCCCAACAGCCCCCTGATTGGGCAG
 AAGCTGAGCGAGATGGACAGCGTGCTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGA
 GACGCCCTTTACCCAATTTGCTGGTTCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACG
 GGGCCTCTCCACCGAGGAGGTGAATACACCTCTGATTTTAATCAGTTCTGCGTTTGAAAGGAAA
 CCCGGTGATATCCGACATCCAAGCACGTCCAATAGACGGATGTGGCTGCCACACTGGCGATAGC
 ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCCTATTCCCAGTTGTGGAAGGAAGAC
 CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCAGCTTAGTAAACTGTTGCAAGAG
 AATGTGCCGTCATATGAAAAAGATCCTGGGTTTGAAGCAGTTTAAATGTCAGAAAGATTGCATGG
 GAACTGGATCAGACTGTACTTGGAGGAAAAGCATTAGAAGTCCTATTCAACCTGGGCTCCAAGG
 TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCAG
 TTCTCACCCCTGCTCCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCA
 CTGTCATCTCCTGGGTTTTCTCTGCTCTTTTATTTGGTGATCCTGGTCTTTTCGGCCGTTTACGT
 CATTTGTGTGCACCTCAGCTGAAAGTTTCGTGCTACTTCTGTGGCTCTCGTGGCTGGCGGCAGGCT
 GCCTTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGGCAGTGCCCTGGAC
 AGGGGGCCTCAGGGAAGGACGTGGAGCAGCCTTATCCCAGGCCTCTGGGTGTCCCAGACAGGTG
 TTCACATCTGTGCTGTCAGGTCAGATGCCTCAGTTCTTGGAAGCTAGGTTCTGCGACTGTTAC
 CAAGGTGATTGTAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGTGTGTGAA
 TCGGACAGCCTCCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCCTGGA
 CACTCAGGAGGGTCAAAGGAGACTTGGTCGCACCACTCATCCTGCCACCCCCAGAATGCATCCT
 GCCTCATCAGGTCCAGATTTCTTTCCAAGGCGGACGTTTTCTGTGGAATTCTTAGTCTTTGGCC
 TCGGACACCTTCATTCGTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCAC
 ACTCAGATCCACAGAGCCCAGGATCAAGGGACCACTGCAGTGGCAGCAGGACTGTTGGGCCCCC
 ACCCCAACCCTGCACAGCCCTCATCCCCCTTTGGCTTGAGCCGTCAGAGGCCCTGTGCTGAGTGT
 CTGACCGAGACACTCACAGCTTTGTCATCAGGGCACAGGCTTCCTCGGAGCCAGGATGATCTGTG
 CCACGCTTGCACCTCGGGCCCATCTGGGCTCATGCTCTCTCTCTGCTATTGAATTAGTACCTAG
 CTGCACACAGTATGTAGTTACCAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEHGAEPPEPSAGASSNWTTLPPPLF
 SKVVIVLIDALRDDFVFGSKGVKFMPTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
 VDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTTSFFVSDYTEVDNNV
 TRHLDKVLKRGDWDILILHYLGLDGHIGHISGPN SPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
 NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
 275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTACTGCTGCTGTTGGAGTACAAC
TCCCTATAGAAAACAACCTGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATCTCAT
AGCAGTTCAGATAAAAACCTACATACGCCCAGAGATCTTCTTTCATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAGTTTTGTCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT
GGCTGCCCCAAAAGGAATCAGCACGCCGGCCCTTCATCTTTTATAGGGCTCAGGTGGGCTCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCCGGATGGTTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAATTTGAGAACAGGAAACACATTGAATTTTCATTTCAACCAGTTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAACTGCCCCATTGAACGCCTTCCTCGCTA
ATTTGAACTAATTGTATAAAAACACCAAACCTGCTCACT }

FIGURE 142

MLLLLLLEYNFPIENNCQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLD SGNLI AVPDKNYIRPEI
FFALASSLSSASA EKGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAAQKESARRPFI
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 144

MLGLPWKGGLSWALLLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGGAAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCCAGGTGGCTGAGAACCGCCCGGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTCACCGGCTGCATCAATGCCA
CCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTTGGTTGGAGAGGGGCGCAGGACT
TCGGGTACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGATCTGGCTCATGGTGAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCCAAACCCACGCGTGTTCTGAAGGTGCCAGGAGCGGCGATGCACTCGCACTGCAAA
TGCCGCTCCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACCTGCGCCACCTGCTTCACAGTACTTCCCAACAACCTCTTAGAGGTAG
GTGTATTCCCGTTTTACAGATAAGGAACTGAGGCCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCTGGCTTGCTAACCCAGGTTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTTAGAAATAAATGAAAACACCTGA

FIGURE 146

MRKHLSSWWLATVCMLLFSHLSAVQTRGIKHKRKNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQ
VLWRLVQELCSLKHCEFWLERGAGLRVTMHQPVLLCLLALIWLMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATTT
 CCTGATGATTTATAGACTCAAAGAAAACTATGTTTCAGAAGCTCTCTTCTTCTTCTGGCCTCCTCT
 CTGTCTTCTTTCCCTCTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
 AAATCTTTCATTTTGCTTGTGAGTGGGGTAGGTCACTGAGTCTTAGTTTTTATTTTTTGAAATTT
 CAACTTTCAGATTCAGGGGGTACATGTGAAGGTTTGTTTTATGAGTATATTGCATTGATGTCTGAGG
 TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLLLISSIYSECKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTACACGAACTTCAGCACCCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
TCCCGCGCGCCCCAACCTTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCCGCCAGTCC
CGGCCCTCTCCCGCCCCACACCACCCTCTGGCTCTTCTGTTTTACTCCTCTTTTCATTCTATA
ACAAAAGCTACAGCTCCAGGAGCCAGCGCCGGGCTGTGACCCAAGCCGAGCGTGAAGAATGGGGTT
CCTCGGGACCGGCACTTGATTCTGGTGTGTAGTGCTCCCGATTCAAGCTTTCCCCAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAAAAAACATATCCTCCAGAAAAACAAGCCAGGTGAGAGCAACTATTCTTT
TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAAGGAAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACCTG
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA
TCAACTAGACGGGACTCCTTTAACCCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTTATGAAG
AAAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAAACTACTTAATCTCGGCCCTTATCACGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCCAATAAGCCCAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAACCTACAGTGAAGACAACCTTTGAGGAACCTCCA
ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAAGAAGCAAAAGAGAAGAAAA
CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAATATGGAACAATATCT
CCAGAAGAAGGTGTTTCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAAACAAGCT
AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGC
CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAAGGGAATAAAGAAGATTATGACCTTT
CAAAGATGAGAGACTTCAATCAATAAAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAAAATGGCAAAAGATCCAGGAGTCTTTCAA
CTGTTTCAGAAAAACATAATATAGCTTAAACACTTCTAATTCTGTGATTAAAATTTTTTGACCCAAGG
GTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGCTTTCTTCCC
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIKEKVTMAAIQDGLAKGENDETVSNTLTLTNGLE
RRTKTYSEDNFEELQYFPNFIYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAGCTCAA
GATGGTCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC
ATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCCGTCATCCTGGGTGTCCAGGGTGGAAAG
CCAGTGCCTGTGCATGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTACCTTCTACCGGCGGGACATGGGGCTCACC
TCCAGCTTCGAGTCGGCTGCCTACCCGGGCTGGTTCCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCAGACTCACCCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCATCACAGACTTCTACTTCC
AGCAGTGTGAC**TAGG**GCAACGTGCCCCCAGAACTCCCTGGGCAGAGCCAGCTCGGGTGAGGGGT
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCACTTTGTCTTCTGGTTCCAGTTTGGATAAATTCTGAGATTTGGAGCTCAGT
CCACGGTCTCCCCCACTGGATGGTGCTACTGCTGTGGAACCTTGTA AAAACCATGTGGGGTAAA
CTGGGAATAACATGAAAAGATTTCTGTGGGGTGGGGTGGGGGAGTGGTGGGAATCATTCCTGCT
TAATGGTAACAGCAAGTGTACCTGAGCCCCGAGGCCAACCATCCCCAGTTGAGCCTTATA
GGGTCAGTAGCTCTCCACATGAAGTCTGTCACTCACCAGTGTGCAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTTGGCCCAGCCCCACCCCTTCCCTTTAATCCTGCCACTGTGCTATA
TGCTACCTTTCTATCTCTTCCCTCATCATCTTGTGTGGGCATGAGGAGGTGGTGTATGTGAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTTAAAAACCCAA
GATACAATCAAAATCCCAGATGCTGGTCTCTATTCCCATGAAAAAGTGCTCATGACATATTGAGA
AGACCTACTTACAAAGTGGCATATATTGCAATTTATTTTAAATTAAGATACCTATTTATATATT
TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTGTAGCAATGTGAGGGTGGTGGCAGTAT
AGGTGATTTTTCTTTAATTCTGTAAATTTATCTGTATTTCTAATTTTTCTACAATGAAGATGA
ATTCTTGTATAAAAAATAAGAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGA
TTGTCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTTGAATCGAGCTCTGCTGCTCTGGTTGG
TTGTAGTAGTGATCAGGAAACAGATCTCAGCAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT
GTGGCTGGAATCTCTGGGTAAGGAACTTAAAGAACAAAAATCATCTGGTAATCTTTCTCTAGAAG
GATCACAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGTGCTGCTGGTTGA
ATTGTGTCCCCCTCAAATTCACATCCTTCTTGGAACTCTAGTCTGTGAGTTTATTTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCTATGCTGGATGAAGGTAGACCTAAATTCATAT
GACTGGTTTCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTTCAGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACC
ATCAGAAGCTTGGAAGAGGCAAAGAAGAAATCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAGAAATAAATTTCCGGCTGTTTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATCCCTGT
CTCCTCGTGTTTACATCTGTGTGTGTCCCCCTCCCAATGTACCAAAGTTGTCTTTGTGACCAA
TAGAATATGGCAGAAAGTGTATGGCATGCCACTTCCAAGATTAGGTTATAAAAGACACTGCAGCTTC
TACTTGAGCCCTCTCTCTGCCCCACCGCCCCCAATCTATCTTGGCTCACTCGCTCTGGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCATATAAGAGACTTACGTGGTAAAAAATGAAGTCCTT
GCCCCAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTTT
AAGTTGCTCAGTTTTGGTCTAACTTGTATGCAGCAATAGATAAATAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEEISVVPNRWLDASLSPVILGVQGS
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGAWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTCTTGG
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACTTCACCCTTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTTC
CAGCCTTATATGCAGGAGGTGGTGGCCCTTCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTTATGTCTCTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTTCCCTGCTAGAAATAA
CAATTAGATGCCCCAAAGCGATTTTTTTTAACCAAAGGAAGATGGGAAGCCAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTTTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTTCATTGTAAGTGGTGTTT
TATACACAGAAAACAATTTATTTTTTAAATAATTGTCTTTTTCCATAAAAAAGATTACTTTCCAT
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTCCATAATCAGTACTTTATATTTATAAA
TGTATTTATTATTATTATAAGACTGCATTTTATTATATCATTTTATTAATATGGATTTATTAT
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT
AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLLLVQGAAAPISSHCRDKSNFQQPYITNRTFMLAKEASL
ADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLPQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCCGACTTGTGACTGAGTGTGCAGTGCCAGCATGTACCAGGTCAGTGCAGAGGGC
TGCCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTCTACAGGTGGTTGCAT
TCTTGGAATGGTCATGGGAACCCACACCTACAGCCACTGGCCCAGCTGCTGCCCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCCTAGAGCCTGCTAG
GCCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACCACGCCCCGT
TGCCCTGTGCCCGCACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCCGGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTTCCTTAGCTTGTGTGTGTGTGCGGGCCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA
ACCACTTGCCATGAAGGGCCAGGATGCCCAGATGCTTGGCCCCTGTGAAGTGTCTGTGAGCAG
CAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAAAACCTGCACCTTCTGCACATTTTGAAAAGAG
CAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGTCTGTCACTTTCTCTCAGGAAAGGTTTTCAAA
GTTCTGCCCATTTCTGGAGGCCACCACTCCTGTCTCTTCTCTTTTCCCATCCCCTGCTACCCTG
GCCCAGCACAGGCACTTTCTAGATATTTCCCCCTTGTCTGGAGAAGAAAGAGCCCCCTGGTTTTATT
TGTTTGTTTACTCATCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTTATCCAAATAAATAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLSLFLQVVAFLAMVMGTHYSHWPSCCPKQDTSELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVS LACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
 GACCGTTCAATGTGGCTCTGAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
 CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAAGTAGTGTGCAACAGGGGACTATTCA
 ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTTGAAGGCCACCAA
 GATTTGTGTGACGGGCAAAGCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAATTACACAGAGG
 CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCTACATCGGCTTCCCTGTA
 GAGCTGAACACAGTCTATTTTCATTGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG
 CCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAA
 AGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTAAGAAGAATGAGGAGACA
 GTAGAAGTGAAGTTCACAACCACTCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC
 TATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
 TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCTACTTGTGGC
 AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGTCCCTTTCCCTCT
 GGATAACAACAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCTGCTGTCTCTGCTGGTGGCCA
 CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT
 TCTACCACCACACTACTGCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA
 TCACACAATTTGTTACTTCACTGAATTTCTTCAAACCATTCGAGAAGTGAGGTCATCCTTGAAA
 AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAGAAGGCA
 GCAGACAAAGTCGTCTTCTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
 GAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTCTTCCCCCTTGCTTTAACCTTTTCTGCA
 GTGATCTAAGAAGCCAGATTCACTGACACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA
 AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT
 CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
 ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPGLRDLRVEPVTTTSVATGDYSILMNVSWV
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPELNTVYFIGAHNIP
NANMNEDGPSMSVNFETSPGCLDHIMKYKKKCVKAGSLWDPNITACKKNEETVEVNFNTTTPLGNYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPCQTGVPFPLDNNK
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGSFSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVKKQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
 GTCGATATTGGGGCTTGCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
 CTTTTTCCAAAAGCCTGAGAGTTGCCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
 ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
 GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCCAGTGTAGGA
 ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCCATCCAGCAA
 GAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTTCCAGTTGGAGAAGGTGCT
 GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTTAAGAGGTGCATATCC
 ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVMSRNIESRSTSPWNYTWTWDPNRYPSEVVQAQCRNLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGC SVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAAGGACTCCCAGG
ACAGAGAGTGCACAACTACCCAGCACAGCCCCCTCCGCCCCCTCTGGAGGCTGAAGAGGGATT
CAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCTTGGGGGGGGGAGCAC
AGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCCTGTGCCCTGGTTCTTGCTGTCT
TGGCACTGGGCGGAGCCCAAGTGGTCTTTCTCTGGAGAGGCTTGTGGGGCTCAGGACGCTACC
CACTGCTCTCCGGGCTCTCTGCGGCTCTGGGACAGTGACATACTCTGCCTGCCTGGGGACAT
CGTGCTGTCTCCGGGCCCCGTGCTGGGCGCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGCCGTGCATGGGCACTGG
GAAGAGCCTGAAGATGAGGAAAAGTTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAA
TGCCTCTCTCCAGGCCAAGTCGTCTCTCTTCCAGGCCTACCTACTGCCCCGTGCGTCTCTGC
TGGAGGTGCAAGTGCCTGTGCCCCCTTGTGCACTTGGTCACTGTGGGCTCTGTGGTATATGAC
TGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTCTTACTCAGGCCAGGTACGAGAA
GGAACCTCAACCACACACAGCAGCTGCCTGCCCTGCCCTGGCTCAACGTGTACGAGATGGTGACA
ACGTGCATCTGGTTCTGAATGTCTCTGAGGAGCAGCACTTCGGCCTCTCCCTGTACTGGAATCAG
GTCCAGGGCCCCCAAAACCCCGTGGCACAAAACCTGACTGGACCGCAGATCATTACCTTGA
CCACACAGACCTGGTTCCCTGCCCTCTGTATTAGGTGTGGCTCTGGAACCTGACTCCGTTAGGA
CGAACATCTGCCCCCTCAGGGAGGACCCCGCGCACACCAGAACCTCTGGCAAGCCCGCCGACTG
CGACTGCTGACCCTGCAGAGCTGGCTGCTGGACGCACCGTGTCTCGCTGCCCGCAGAAGCGGCACT
GTGCTGGCGGGCTCCGGGTGGGGACCCCTGCCAGCCACTGGTCCCACCGCTTTCTGGGAGAACG
TCACTGTGGACAAGGTTCTCGAGTTCCCATTTGCTGAAAGGCCACCCTAACCTCTGTGTTCAAGTG
AACAGCTCGGAGAAGCTGCAGCTGCAGGAGTGCTTGTGGGCTGACTCCCTGGGGCTCTCAAAGA
CGATGTGCTACTGTTGGAGACACGAGGCCCCAGGACAACAGATCCCTCTGTGCCTTGAACCCA
GTGGCTGTACTTCACTACCCAGCAAAGCCTCCACGAGGGCAGCTCGCCTTGGAGAGTACTTACTA
CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCTG
CCCCATGGACAAATACATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCCTACTCTTTGCCG
CTGCGCTTTCCCTCATCTCTCTTCAAAAAGGATCACGCGAAAGGGTGGCTGAGGCTCTTGAAA
CAGGACGTCGCTCGGGGGCGGCCGCCAGGGGCCGCGCGCTCTGCTCTCTACTCAGCCGATGA
CTCGGGTTTCGAGCGCTGGTGGGCGCCCTGGCGTCCGGCCTGTGCCAGCTGCCGCTGCGCGTGG
CCGTAGACCTGTGGAGCCGTCTGTAAGTGTGAGCGCGCAGGGGCCCGTGGCTTGGTTTACGCGCAG
CGGCGCCAGACCCTGCAGGAGGGCGCGTGGTGGTCTTGCTCTTCTCTCCCGTGGCGTGGCGCT
GTGCAGCGAGTGGCTACAGGATGGGGTGTCCGGGCCCCGGGCGCACGGCCCGCACGACGCTTCC
GCGCTCGCTCAGCTGCGTGTGCGCGACTTCTTGAGGGCCGGCGCCCCGCGAGCTACGTGGGG
GCCTGCTTCGACAGGCTGTCTCCACCCGACGCCGTACCCGCCCTTTCCGACCGTGCCGCTCTT
CACACTGCCCCCAACTGCCAGACTTCTTGGGGGCCCTGCAGCAGCCTCGCGCCCCGCTTCCG
GGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCCGGGCCCTTACGCCAGCCCTGGATAGCTACTTC
CATCCCCGGGGACTCCCGCGCGGGACGCGGGGTGGGACCAGGGGCGGGACCTGGGGCGGGGA
CGGGACTTAAATAAAGGCAGACGCTGTTTTCTAAAAAA

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT
GCTCACGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCAGTCCAGCAACTTTGA
AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGGACTGGGTGGCAAAGAAGGGCTGTCAAGCGGATCACCCGGAAGTCTGCG
AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT
GTCAGTGCGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTGAGCTCTCTGCAGCACACTAC
CCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTGATGATTGTTTCATCCTA
CCCCACGCCAATCCGTGCAGGCGATGGCCACCGGCTAACCTTGGGAAGACATCTTCCATGACCTG
TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGAGGGAAGCAGAGAGA
ATATGAGTTCTTCGGCCTGACCCCTGACACAGAGTTCTTGGCACCATCATGATTGCGTTCCCA
CCTGGGCCAAGGAGAGTGCCCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC
TACTCCTTCTCCGGAGCCTTCCGTGTCTCCATGGGCTTCCCTCGTCGAGTACTCTGCTACCTGAG
CTACAGATATGTACCAAGCGCCTGCACCTCCCAACTCCCTGAACGTCCAGCGAGTCTGACTT
TCCAGCGCTGCGCTTCATCCAGGAGCAGTCCCTGATCCCTGTCTTGACCTCAGCGGCCCCAGC
AGTCTGGCCCCAGCCTGTCCAGTACTCCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC
TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTATGCCCCAACGCTGCCCCGTGAG
GTCGGGCCCCCATCTATGCACCTCAGGTGACCCCGAAGCTCAATTCCTATTCTACGCCCCACA
GGCCATCTCTAAGGTCCAGCCTTCTCTCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
CCTATGGGGTATGCATGGAAGGTTCTGGCAAAGACTCCCCACTGGGACACTTTCTAGTCTTAA
CACCTTAGGCCTAAAGGTGAGCTTCCAGAAAGAGCCAGCAGCTGGAAGCTGCATGTTAGGTGGCCT
TTCTCTGCAGGAGGTGACCTCCTTGGCTATGGAGGAATCCCAAGAAGCAAAATCATTGCACCAGC
CCCTGGGGATTTGCACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGGACA
CCACAGTACCTAAAGGGCCAGCTCCCCCTCCTCTCCTCAGTCCAGATCGAGGGCCACCCATGTC
CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
TGCTGGAGTCCCTTGTGTGTCCCAAGGATGAAGCCAAGAGCCCAGCCCCCTGAGACCTCAGACCTG
GAGCAGCCACAGAACTGGATTCTCTTTTTCAGAGGCTGGCCCTGACTGTGCAGTGGGAGTCCCTG
AGGGGAATGGGAAGGCTTGGTGCTTCCCTGTCCTACCCAGTGTACATCCTTGGCTGTCA
ATCCCATGCCTGCCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC
AGAGGGAGTGGCATGCAGGGCCCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA
AGGACTGCAGCGGGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGCGTGCTCGCTGAGCCCTG
CAAGGCAGAAATGACAGTGCAAGGAGGAAATGCAGGGAACTCCCGAGGTCCAGAGCCCCACCTC
CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCCCTCTCCTTGGCCATTCTTGCCAGTTTC
ACAATCTAGCTCGACAGAGCATGAGGCCCTGCCCTTCTGTGCTATTGTTCAAAGGTGGGAAGAGA
GCCTGGAAAAGAACCAGGCCTGGAAGAAGACAGAGGAGGCTGGGCAGAACCAGAACACCTGC
ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGCTCA
TTCCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTGAGCTTCATTCCTCTGATAGAACAAAGC
GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTTCTGCAGGCAGGAGTTTCAGACCCT
ATCCTGAGAATGGGGTTTGAAGGAAGGTGAGGGCTGTGGCCCCCTGGACGGGTACAATAACACAC
TGTAAGTGTGACAACTTTGCAAGCTCTGCCTTGGGTTGAGCCCATCTGGGCTCAAATTCAGC
CTCACCCTCACAAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTTCCCTC
ATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
TCTTTAAAGTGCTTAATAGTGCCTGGTACATGGGCAGTGCCCAATAACGGTAGCTATTTAAAAA
AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
 VAKKGCQRITRKSCNLTVETGNLTelyyARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCIS
 KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHGKGQREYEFFGLTPDT
 EFLGTIMICVPTWAKESAPYMCrvKTLpDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
 PNSLNVQRVLTFQPLRFIQEHVLIpVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
 YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPPSY
 APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPGQLQKEPPAGSCMLGGLSLQEVTSLAM
 EESQEAKSLHQPLGICTDRTSDPNVLHSGEETPQYLKGQLPLLSSVQIEGHPMSLPLQPPSGPC
 SPSDQGPSPWGLLESVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAGTCACCGGGCCCGCGGTGGCCACAACATGG
 CTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTTCGTGCTGGGGGCGCTCTGGTGGGTCCCGGGCCAG
 TCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
 GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTTA
 AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAACCTTTGGGCTGGA
 AGTGTGAACACAGTTTGGATATTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA
 AGAAGAGCTACATATTCCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATT
 TTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAACCTGGAGGACTCTGTACCTGAAGAG
 TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
 ACTTGACCTGTGCCTGAGCCCGAGGCATTAGAGCTGATTAGAGGATGGAGAAGGTGCTTTCT
 CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGT
 CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTTGAAGAAATTCTGCACGATAA
 ATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
 AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
 CATTACAGCAAAGGATTCGTTGGCATCAAAATCTAAGTTTGTTTTACAAAGATTGTTTTTAGTA
 CTAAGCTGCCTTGGCAGTTTGCATTTTGTAGCCAAACAAAATATATTATTTCCCTTCTAAGTA
 AAAAAAAAAAAAAAAAAA

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCAGCTTGTCTAGAGGCCGGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCCTTAGGACATACACGGGACCCCTAAGCTTC
AGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCGGGCACAGG
CGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTCCGGGCAGCTGGGCTCGGGC
GGCGGGAGTAGGGCCCGCAGGGAGGCGAGGAGGCTGCATATTCAGAGTCCGGGGCTGCGCCCTG
GGCAGAGGCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGCGTGG
TCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCGCGCCGCTGGTCTAGC
GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAAGT
GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCCTGGCTTGTGAGAGTGAGGGAGGAGTCCCTCC'TCA
GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG
ACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTGGAGGAATGGAGATGGGCAACATCTGG
TGCCTGCCAGATCTCTACCAGTGGTCTGATGGAAGCAATTCAGTACCGAACTGGTACACAG
ATGAACCTTCCCTGCGGAAGTGAAAGTGTGTTGTGATGTATCACAACCAACTGCCAATCCTGGC
CTTGGGGTCCCTACCTTTACCAGTGAATGATGACAGGTGTAACATGAAGCACAATTATATTTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCCGTAGAAAAGCCTTATCTTACAAATCAACCAG
GAGACACCCATCAGAATGTGGTTGTACTGAAGCAGGTATAATTCCCAATCTAATTTATGTTGTT
ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTGGAACTGTTGTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAACTAGTCCAAACAGTCTACACTGTGGATTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTATAATAACTCATTGACTTGGTCCAGAAATTTGTAATCT
GGATCTGTATAAGGAATGGCATCAGAACAAATAGCTTGGAAATGGCTTGAAATCACAAGGATCTGC
AAGATGAAGTGAAGCTCCCCCTTGAGGCAATATTAAAGTAATTTTATATGTCTATTATTTCA
TTTAAAGAAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAA
ACTTCAAACCTTCAAGCAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA
TGTGTGTAGAAAGCAATTCCTTTTATTTCTTTACCTTTTATAAGTTGTTATCTAGTCAATGTAA
TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGATAAAA
ATGAAGTGTCTAATATTTATTTTATGGCATCTCATTTTCAATACATGCTCTTTTGATTAAAG
AACTTATTAAGTGTGCAACTGAATTCACACACACAAATATAGTACCATAGAAAAAGTTTGT
TTTCTCGAAATAATTCATCTTTACGCTTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGA
AATAAGAAGCTATTTTCAATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT
TGTCTAATTTCAATTTGCAAGACATGTGCCTTATAATATTTTATAGCTTAAATTAACAGATT
TTGTAATAATGTAACTTGTGTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAAG
TGACATACACAATATAAATCATATGTCTTACACGTTGCCTATATAATGAGAAGCAGCTCTCTGA
GGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTCCGGGTTTGGG
ATTGACACTGGAGGCAGATAGTTGCAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTG
ACTATATTAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG
ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAACACGAAATCGTGTGAAATGGGTTGG
AACCCATCAGTGATCGCATATTCATTGATGAGGTTTGGCTTGAGATAGAAAATGGTGGCTCCTTT
CTGTCTTATCTCCTAGTTCTTCAATGCTTACGCTTGTCTTCTCAAGAGAAAGTTGTAAGTCT
CTGGTCTTATATGTCCCTGTGCTCCTTTTAAACCAATAAAGAGTTCTTGTCTTCTGGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFRCRRVVGQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDGFWIGLWRNGDGQTSACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSkgRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217